On the Ingleton-Violations in Finite Groups

Wei Mao, Matthew Thill, and Babak Hassibi, *Member, IEEE*

Abstract

Given n discrete random variables, its entropy vector is the $2^n - 1$ dimensional vector obtained from the joint entropies of all non-empty subsets of the random variables. It is well known that there is a one-to-one correspondence between such an entropy vector and a certain group-characterizable vector obtained from a finite group and n of its subgroups [3]. This correspondence may be useful for characterizing the space of entropic vectors and for designing network codes. If one restricts attention to abelian groups then not all entropy vectors can be obtained. This is an explanation for the fact shown by Dougherty et al [4] that linear network codes cannot achieve capacity in general network coding problems (since linear network codes form an abelian group). All abelian group-characterizable vectors, and by fiat all entropy vectors generated by linear network codes, satisfy a linear inequality called the Ingleton inequality. It is therefore of interest to identify groups that violate the Ingleton inequality. In this paper, we study the problem of finding nonabelian finite groups that yield characterizable vectors which violate the Ingleton inequality. Using a refined computer search, we find the symmetric group S_5 to be the smallest group that violates the Ingleton inequality. Careful study of the structure of this group, and its subgroups, reveals that it belongs to the Ingleton-violating family $PGL(2, q)$ with a prime power $q \geq 5$, i.e., the projective group of 2×2 nonsingular matrices with entries in $\mathbb{F}_q$. We further interpret this family of groups, and their subgroups, using the theory of group actions and identify the subgroups as certain stabilizers. We also extend the construction to more general groups such as $PGL(n, q)$ and $GL(n, q)$. The families of groups identified here are therefore good candidates for constructing network codes more powerful than linear network codes, and we discuss some considerations for constructing such group network codes.

Index Terms

Portions of this work were presented at the Forty-Seventh Annual Allerton Conference on Communication, Control, and Computing, 2009 [1] and the 2010 IEEE International Symposium on Information Theory [2]. The authors are with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125 USA (email: wmao@caltech.edu, mthill@caltech.edu, hassibi@caltech.edu). This work was supported in part by the National Science Foundation under grants CCF-0729203, CNS-0932428 and CCF-1018927, by the Office of Naval Research under the MURI grant N00014-08-1-0747, and by Caltech's Lee Center for Advanced Networking.

Finite groups, entropy vectors, Ingleton inequality, network coding, group network codes.

I. Introduction

Let $\mathcal{N} = \{1, 2, \ldots, n\}$, and let $X_1, X_2, \ldots, X_n$ be n jointly distributed discrete random variables. For any nonempty set $\alpha \subseteq \mathcal{N}$, let X_α denote the collection of random variables $\{X_i : i \in \alpha\}$, with joint entropy $h_\alpha \triangleq H(X_\alpha) = H(X_i; i \in \alpha)$. We call the ordered real $(2^n - 1)$-tuple $(h_\alpha : \emptyset \neq \alpha \subseteq \mathcal{N}) \in \mathbb{R}^{2^n-1}$ an *entropy vector*. The set of all entropy vectors derived from n jointly distributed discrete random variables is denoted by Γ_n^*. It is not too difficult to show that the closure of this set, i.e., $\overline{\Gamma_n^*}$, is a *convex cone* [5].

The set $\overline{\Gamma_n^*}$ figures prominently in information theory since it describes the possible values that the joint entropies of a collection of n discrete random variables can obtain. From a practical point of view, it is of importance since it can be shown that the capacity region of any arbitrary multi-source multi-sink *wired* network, whose graph is acyclic and whose links are discrete memoryless channels, can be obtained by optimizing a linear function of the entropy vector over the convex cone $\overline{\Gamma_n^*}$ and a set of linear constraints (defined by the network) [6], [7]. Despite this importance, the entropy region $\overline{\Gamma_n^*}$ is only known for $n = 2, 3$ random variables and remains unknown for $n \geq 4$ random variables. Nonetheless, there are important connections known between $\overline{\Gamma_n^*}$ and matroid theory (since entropy is a submodular[1] function.) [8], determinantal inequalities (through the connection with Gaussian random variables) [9], and quasi-uniform arrays [10]. However, perhaps most intriguing is the connection to finite groups which we briefly elaborate below.

A. Groups and Entropy

Throughout this paper we use group theory notation defined in Section II. Let G be a finite group, and let $G_1, G_2, \ldots, G_n$ be n of its subgroups. For any nonempty set $\alpha \subseteq \mathcal{N}$, the group $G_\alpha \triangleq \bigcap_{i\in\alpha} G_i$ is a subgroup of G. Define $g_\alpha = \log \frac{|G|}{|G_\alpha|}$. We call the ordered real $(2^n - 1)$-tuple $(g_\alpha : \emptyset \neq \alpha \subseteq \mathcal{N}) \in \mathbb{R}^{2^n-1}$ a (finite) *group characterizable vector*. Let Υ_n be the set of all group characterizable vectors derived from n subgroups of a finite group.

The major result shown by Chan and Yeung in [3] is that $\overline{\Gamma_n^*} = \overline{\mathrm{cone}(\Upsilon_n)}$, i.e., the closure of Γ_n^* is the same as the closure of the cone generated by Υ_n. Specifically, every group characterizable vector is

[1] A set function f defined on the subsets of $\mathcal{N}$ is *submodular* iff $f_\alpha + f_\beta - f_{\alpha\cap\beta} - f_{\alpha\cup\beta} \geq 0$ for all $\alpha, \beta \subseteq \mathcal{N}$.

an entropy vector, whereas every entropy vector is arbitrarily close to a scaled version of some group characterizable vector.

To show the first part of this statement, let Λ be a random variable uniformly distributed on the elements of G and define $X_i = \Lambda G_i$ (the left coset of G_i in G with representative Λ) for $i = 1, \ldots, n$. Then X_i is uniformly distributed on G/G_i and $H(X_i) = \log \frac{|G|}{|G_i|}$. To calculate the joint entropy $h_\alpha = H(X_\alpha)$ for a nonempty subset $\alpha \subseteq \mathcal{N}$, let $\mathcal{X}_\alpha$ denotes the set of all coset tuples $\{(xG_i : i \in \alpha) \mid x \in G\}$. Consider the intersection mapping $\Theta_\alpha : \mathcal{X}_\alpha \to G/G_\alpha$, where for all $x \in G$,

$$\Theta_\alpha : (xG_i : i \in \alpha) \mapsto \bigcap_{i \in \alpha} xG_i = xG_\alpha. \tag{1}$$

Θ_α is a well defined onto function on $\mathcal{X}_\alpha$, and it is one-to-one since if $(xG_i : i \in \alpha)$ and $(x'G_i : i \in \alpha)$ are mapped to the same coset $xG_\alpha = x'G_\alpha$, then $x^{-1}x' \in G_\alpha$ and so $x^{-1}x' \in G_i$ for all i, which implies $(xG_i : i \in \alpha) = (x'G_i : i \in \alpha)$. So $H(X_\alpha) = H(\Theta_\alpha(X_\alpha))$, and as $\Theta_\alpha(X_\alpha) = \Lambda G_\alpha$, we have

$$h_\alpha = H(\Lambda G_\alpha)) = \log \frac{|G|}{|G_\alpha|} = g_\alpha.$$

Thus indeed every group-characterizable vector is an entropy vector. Showing the other direction, i.e., that every entropy vector can be approximated by a scaled group-characterizable vector is more tricky (the interested reader may consult [3] for the details). Here we shall briefly describe the intuition.

Consider a random variable X_1 with alphabet size N and probability mass function $\{p_i, i = 1, \ldots, N\}$. Now if we make T copies of this random variable to make sequences of length T, the entropy of X_1 is roughly equal to the logarithm of the number of strongly typical sequences, divided by T. These are sequences where X_1 takes its first value roughly Tp_1 times, its second value roughly Tp_2 times and so on. Therefore assuming that T is large enough so that the Tp_i are close to integers (otherwise, we have to round things) we may roughly write

$$H(X_1) \approx \frac{1}{T} \log \begin{pmatrix} & & T & & \\ Tp_1 & Tp_2 & \ldots & Tp_{N-1} & Tp_N \end{pmatrix},$$

where the argument inside the $\log$ is the usual multinomial coefficient. Written in terms of factorials this is

$$H(X_1) \approx \frac{1}{T} \log \frac{T!}{(Tp_1)!(Tp_2)! \ldots (Tp_N)!}. \tag{2}$$

If we consider the group G to be the symmetric group S_T, i.e., the group of permutations among T objects, then clearly $|G| = T!$. Now partition the T objects into N sets each with Tp_1 to Tp_N elements, respectively, and define the group G_1 to be the subgroup of S_T that permutes these objects *while respecting the partition*. Clearly, $|G_1| = (Tp_1)!(Tp_2)! \ldots (Tp_N)!$, which is the denominator in (2).

Thus, $H(X_1) \approx \frac{1}{T} \log \frac{|G|}{|G_1|}$, so that the entropy $h_{\{1\}}$ is a scaled version of the group-characterizable $g_{\{1\}}$. This argument can be made more precise and can be extended to n random variables—see [3] for the details. We note, in passing, that this construction often needs T to be very large, so that the group G and the subgroups G_i are huge.

B. The Ingleton Inequality

As mentioned earlier, entropy satisfies submodularity and is connected to the notion of matroids. A matroid is defined by a ground set S and a rank function r (written as $r(\{\cdot\})$ or $r_{\{\cdot\}}$) defined over subsets of S, that satisfies the following axioms:

1) r is always a non-negative integer, and $r(U) \leq |U|$, $\forall U \subseteq S$.
2) r is monotonic: if $U \subseteq W \subseteq S$, then $r(U) \leq r(W)$.
3) r is submodular.

Axioms 2) and 3), together with positiveness, are called the *Shannon inequalities* for a set function. A matroid is defined in a way to extend the notion of a collection of vectors (in some vector space) along with the usual definition of the rank. It is called *representable* if its ground set can be represented as a collection of vectors (defined over some finite field) along with the usual rank function. Determining whether a matroid is representable or not is, in general, an open problem.

In 1971 Ingleton showed that for $n = 4$, the rank function r of any representable matriod must satisfy the inequality [11]

$$r_{12} + r_{13} + r_{14} + r_{23} + r_{24} \geq r_1 + r_2 + r_{34} + r_{123} + r_{124}$$

(where for simplicity we write r_{ij} and r_{ijk} for $r_{\{i,j\}}$ and $r_{\{i,j,k\}}$, respectively). In fact, these *Ingleton inequalities*, together with the Shannon inequalities and their combinations, are the only inequalities the rank function of a representable matroid needs to satisfy (which are called linear rank inequalities) when $n = 4$ (see [12]). Furthermore, [12] shows that the rank function of any representable matroid is necessarily an entropy vector, but not every linear rank inequality is respected by a general entropy vector. For example, there are entropy vectors that violate the Ingleton inequality (e.g. [12], [13]), so that entropy is generally not a representable matroid. Using non-representable matroids, [4] constructs network coding problems that cannot be solved by linear network codes (since linear network codes are, by definition, representable).

When $n \geq 5$, there are many more linear rank inequalities besides the Shannon ones. But since the focus of this paper is the simplest case $n = 4$ with only one such inequality, we refer the interested readers

to the works of Kinser [14], Dougherty *et al*. [15]–[17] and Chan *et al*. [18] for recent development in this area.

From this point on we shall only study the Ingleton inequality, with $n = 4$. In the case of entropy vectors, it is written as

$$h_{12} + h_{13} + h_{14} + h_{23} + h_{24} \geq h_1 + h_2 + h_{34} + h_{123} + h_{124}. \tag{3}$$

The following sufficient condition is proposed in [12] for four general random variables X_1, X_2, X_3 and X_4 to satisfy (3):

Lemma 1: If there exists a random variable Z that is a *common information* for X_1 and X_2, i.e., $H(Z|X_1) = H(Z|X_2) = 0$ while $H(Z) = I(X_1; X_2)$, then (3) is satisfied.

In general common information does not exist for two arbitrary random variables, but when the entropies correspond to ranks of vector subspaces, their common information does exist [12] and that is why representable matroids respect Ingleton. In Section III we will prove a similar condition for groups to satisfy Ingleton, by constructing a common information.

As $\overline{\Gamma_n^*} = \overline{\mathrm{cone}(\Upsilon_n)}$, we know there must exist finite groups, and corresponding subgroups, such that their induced group-characterizable vectors violate the Ingleton inequality. In [19] it was shown that abelian groups cannot violate the Ingleton inequality, thereby giving an alternative proof as to why linear network codes (and even the more general abelian group network codes (defined below)) cannot achieve capacity on arbitrary networks, as the underlying groups for linear network codes are abelian. So we need to focus on non-abelian groups and their connections to nonlinear codes. Note that in the context of finite groups, the Ingleton inequality can be rewritten as

$$|G_1||G_2||G_{34}||G_{123}||G_{124}| \geq |G_{12}||G_{13}||G_{14}||G_{23}||G_{24}|. \tag{4}$$

C. Group Network Codes

A communication network is usually represented by a directed acyclic graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where the node set $\mathcal{V}$ and the edge set $\mathcal{E}$ model the communication nodes and channels respectively. Let $\mathcal{S} \subset \mathcal{V}$ be the set of source nodes and $\mathcal{D}(s)$ be the set of sink nodes demanding source s for each $s \in \mathcal{S}$. For any node v and any edge e, $\mathcal{I}(v)$ and $\mathcal{I}(e)$ denote the sets of incoming edges to v and to the tail node of e, respectively.

A network code should include

1) the assignment of a symbol Y_s from some alphabet $\mathcal{Y}_s$ for a source message at each source s;

2) the encoding of a symbol Y_e in some alphabet $\mathcal{Y}_e$ at each edge e, from the symbols on $\mathcal{I}(e)$. Namely, $Y_e = \phi_e\left(Y_f : f \in \mathcal{I}(e)\right)$ for some deterministic encoding function ϕ_e;
3) the decoding of the symbol Y_s at each $u \in \mathcal{D}(s)$ for all sources s, i.e. Y_s is uniquely determined from the symbols on $\mathcal{I}(u)$: $Y_s = \phi_{u,s}(Y_f : f \in \mathcal{I}(u))$ for some decoding function $\phi_{u,s}$.

It is clear that at each edge e the symbol Y_e is a deterministic function of the source symbols $\{Y_s : s \in \mathcal{S}\}$, which is denoted by φ_e and is called the *global mapping* at e. Also the source random variables $\{Y_s : s \in \mathcal{S}\}$ are usually assumed to be independent and uniform on their respective alphabets.

For example, a linear network code is defined as follows: 1) for each $t \in \mathcal{S} \cup \mathcal{E}$, the alphabet $\mathcal{Y}_t$ is a vector space F^{d_t} over a finite field F with some finite dimension d_t; 2) all encoding/decoding functions are linear: if t is an edge or a sink node, then the encoding/decoding function ϕ_t at t can be written as

$$\phi_t\left(Y_f : f \in \mathcal{I}(t)\right) = \sum_{f \in \mathcal{I}(t)} M_{t,f} Y_f$$

for some matrices $M_{t,f} \in F^{d_t} \times F^{d_f}$. Thus the global mappings at the edges are also linear.

Group network codes were first proposed by Chan in [20], [21], considering the fact that finite groups can generate the whole entropy region, and noting that linear network codes are included as a special case. Suppose G is a finite group, $\{G_e : e \in \mathcal{E}\}$ and $\{G_s : s \in \mathcal{S}\}$ are some of its subgroups. One can construct a network code with $\mathcal{Y}_t = G/G_t$ for each $t \in \mathcal{S} \cup \mathcal{E}$ if the following requirements are met:

(R1) *Source independence*: $H(Y_{\mathcal{S}}) = \sum_{s \in \mathcal{S}} H(Y_s)$, which means that the cardinalities of $G/G_{\mathcal{S}}$ and $\prod_{s \in \mathcal{S}} \mathcal{Y}_s$ (the Cartesian product of the source alphabets) are equal, where $G_{\mathcal{S}} \triangleq \bigcap_{s \in \mathcal{S}} G_s$. This is equivalent to $\prod_{s \in \mathcal{S}} |G_s| = |G|^{|\mathcal{S}|-1} |G_{\mathcal{S}}|$.
(R2) *Encoding*: $\forall e \in \mathcal{E}$, $\bigcap_{f \in \mathcal{I}(e)} G_f \leq G_e$.
(R3) *Decoding*: $\forall s \in \mathcal{S}$, $\bigcap_{f \in \mathcal{I}(u)} G_f \leq G_s$ for each $u \in \mathcal{D}(s)$.

Moreover, the entropy vector for the network symbols $\{Y_t : t \in \mathcal{S} \cup \mathcal{E}\}$ is characterizable by the group G and its subgroups $\{G_t : t \in \mathcal{S} \cup \mathcal{E}\}$.

In Section VIII we discuss some important considerations necessary when constructing group network codes, such as how to ensure the source independence requirement (R1) above. Appendix A provides further detailed discussions of group network codes, including the encoding/decoding construction, the induced entropy vectors, as well as the inclusion of linear network codes. We remark that any group network code constructed from an Ingleton-violating group induces entropy vectors that violate the Ingleton inequality, so potentially they are more powerful than linear network codes. As we shall see further below, this is certainly true of the group network codes that can be obtained from the Ingleton-

violating families in this paper—the PGL and GL groups, especially since both contain linear network codes as subgroups.

D. Discussion

Since we know of distributions whose entropy vector violates the Ingleton inequality, we can, in principle, construct finite groups whose group-characterizable vectors violate Ingleton. Two such distributions are Example 1 in [13], where the underlying distribution is uniform over 7 points and the random variables correspond to different partitions of these seven points, and the example on page 1445 of [22], constructed from finite projective geometry and where the underlying distribution is uniform over $12 \times 13 = 156$ points. Unfortunately, constructing groups and subgroups for these distributions using the recipe of Section I-A results in $T = 29 \times 7 = 203$ and $T = 23 \times 156 = 3588$, which results in groups of size $203!$ and $3588!$, which are too huge to give us any insight whatsoever.

These discussions lead us to the following questions.

1) Could the connection between entropy and groups be a red herring? Are the interesting groups too large to give any insight into the problem (e.g., the conditions for the Ingleton inequality to be violated)?
2) What is the smallest group with subgroups that violates the Ingleton inequality? Does it have any special structure?
3) Can one construct network codes from such Ingleton-violating groups?

In this paper we address the first two questions, and try to lay some groundwork for answering the third. We identify the smallest group that violates the Ingleton inequality—it is the symmetric group S_5, with 120 elements. Through a thorough investigation of the structure of its subgroups we conclude that it belongs to the family of groups $PGL(2, q)$, with $q \geq 5$ being a power of a prime. ($PGL(2, 5)$ is isomorphic to S_5.) We therefore believe that the connection to groups is not a red herring and that there may be some benefit to it.

Having a "recipe" for Ingleton violations, we generalize the family in two directions. Since $PGL(2, q)$ is the quotient group of $GL(2, q)$ modulo the scalar matrices, we explore the subgroups in $GL(2, q)$ and discover several new families of Ingleton violations. On the other hand, the projective general linear group $PGL(n, q)$ can be viewed as the image of a permutation representation induced by the action of the general linear group $GL(n, q)$ on its projective geometry. It turns out that in this context, the Ingleton-violating subgroups of the family $PGL(2, q)$ all have nice interpretations: each of them is the stabilizer for a set of points in the projective geometry. Based on this viewpoint we obtain a few new

families of Ingleton violations, including the groups $PGL(n,q)$, $GL(n,q)$, and further give an abstract construction in general 2-transitive groups.

As mentioned in Section I-C we can use these Ingleton-violating groups to contruct network codes, which have the potential of performing better than linear network codes. However, designing the subgroups for a desirable code is not a trivial task, for example we need to satisfy (R1)–(R3) of the previous subsection. We study the source independence requirement for the subgroups, and give some directions on how to construct them.

Before we proceed to present the details of our results, we would like to mention some recent developments after our first paper [1] on this subject. In [23], Boston and Nan mainly study symmetric groups and discover many new Ingleton violations in the related groups. Furthermore, using the same group action theoretic approach as above (specifically, designing the subgroups to be the stabilizers of certain sets of points[2]), they systematically construct subgroups of a symmetric group to violate Ingleton. Many of these new violations are quite effective (see Section V-C for more discussions). Also, while all the Ingleton-violating groups in this paper are non-solvable, [23] shows that there do exist solvable groups that violate Ingleton. Paajanen [24], however, focuses on the subclasses p-groups and nilpotent groups and shows that with some technical conditions they satisfy Ingleton. Recall that we have the hierarchy of finite groups

Cyclic groups $\subset$ **Abelian groups** $\subset$ **Nilpotent groups** $\subset$ **Solvable groups** $\subset$ **All groups**

and that every nilpotent group is a direct product of groups, each of which is a p-group for a distinct p. Now roughly speaking, we have a guideline for what class of groups one needs to explore to violate Ingleton. For linear rank inequalities in higher dimensions, [25] considers the case $n = 5$ and obtains some results on the groups that satisfy/violate some of these inequalities.

The rest of the paper is organized as follows. Section II provides necessary notations. Section III describes the computer search process of Ingleton-violating groups and proves several conditions that help pruning the search. Having found the smallest violation instance, Section IV studies its structure using group presentations. Section V then generalizes the instance to an Ingleton-violating family in $PGL(2,p)$, and then to $PGL(2,q)$, through explicitly constructing the subgroups in the format of matrices. Furthermore, the preimage group $GL(2,q)$ is also examined and 15 new families of Ingleton

[2]In fact, in the original paper of Chan and Yeung [3] the same type of subgroups are also used in to show that every entropy vector can be approximated by a scaled group-characterizable vector.

violating subgroups are identified, in Section VI. The original family has a deep relation to the theory of group actions, as disclosed in the more abstract Section VII, which leads to several new violation constructions in this framework. Section VIII, however, considers using these groups to build group network codes and obtains some results in that regard. Section IX concludes this paper.

II. NOTATION

We use the following abstract algebra notations. These are fairly standard (and follow Dummitt and Foote [26]). The interested reader, who may not be familiar with all the concepts below, may refer to [26], or any other standard abstract algebra textbook.

$\|G\|$	the order (cardinality) of the set/group G.
$\|g\|$	the order of element g = smallest positive integer m s.t. $g^m = 1$.
x^g	the conjugate of element x by element g in G: $x^g = g^{-1}xg$. (No confusion with the powers of x as g is an element of G.)
X^g	the conjugate of subset X by element g in G: $X^g = \{x^g : x \in X\}$.
$G \cong H$	the group G is isomorphic to the group H.
$H \leq G$	H is a subgroup of G.
$H \trianglelefteq G$	H is a normal subgroup of G, i.e., $H^g = H$, $\forall g \in G$.
gH	the left coset of the subgroup H in G with representative g.
G/H	the set of all left cosets of subgroup H in G. When $H \trianglelefteq G$, G/H is a group, called the factor group or quotient group.
HK or $H \cdot K$	the "set product" of $H, K \subseteq G$: $HK = \{hk : h \in H, k \in K\}$.
$H \times K$	the direct product of groups H and K. The elements are the pairs $\{(h, k) : h \in H, k \in K\}$ and $(h_1, k_1)(h_2, k_2) = (h_1h_2, k_1k_2)$.
G^n	the direct product of n copies of the group G.
$H \rtimes K$	the semidirect product of groups H and K. The elements are the same as $H \times K$, but $(h_1, k_1) \cdot (h_2, k_2) = (h_1 \cdot \varphi(k_1)(h_2), k_1k_2)$ where φ is a homomorphism of K into the automorphism group of H.
$\langle g_1, \ldots, g_m \rangle$, $\langle S \rangle$	the group generated by the elements $g_1, \ldots, g_m$, and by the set S.
$G = \langle S \mid R \rangle$	$\langle S \mid R \rangle$ is a presentation of G. S is a set of generators of G, while R is a set of relations G should satisfy. See Definition 1.
1	the natural number "1", identity element of a group, or the trivial group. The meaning should be clear in different contexts with no confusion.

$\mathbb{Z}_n$	the integers modulo $n \cong$ the cyclic group of order n.
S_n	the symmetric group of degree n, consisting of all permutations on n points.
D_{2n}	the dihedral group of order $2n$.
$\mathbb{F}_q$	the finite field of q elements.
$\mathbb{Z}_n^\times$, $\mathbb{F}_q^\times$	the multiplicative group of units of $\mathbb{Z}_n$, and of $\mathbb{F}_q$, both consisting of all invertible elements under multiplication. $\mathbb{F}_q^\times$ = all nonzero elements of $\mathbb{F}_q$.
$GL(n,q)$	the general linear group of degree n on $\mathbb{F}_q$, which consists of all invertible $n \times n$ matrices with entries from $\mathbb{F}_q$. The identity element for $GL(n,q)$ is usually denoted by I = identity matrix. $\lvert GL(n,q)\rvert = (q^n-1)(q^n-q)(q^n-q^2)\cdots(q^n-q^{n-1})$.
V_q	the center of $GL(n,q)$, consisting of the collection of matrices that commute with every matrix in $GL(n,q)$ = all nonzero scalar matrices = $\{\alpha I : \alpha \in \mathbb{F}_q^\times\}$.
$PGL(n,q)$	the projective general linear group = $GL(n,q)/V_q$. $\lvert PGL(n,q)\rvert = \lvert GL(n,q)\rvert/\lvert V_q\rvert = \lvert GL(n,q)\rvert/(q-1)$. In other words, it is the group of all invertible $n \times n$ matrices with entries from $\mathbb{F}_q$, where matrices that are proportional are considered the same group element.
$SL(2,q)$	the special linear group = all matrices in $GL(2,q)$ with determinant 1. $\lvert SL(2,q)\rvert = \lvert PGL(2,q)\rvert$.
$PSL(2,q)$	the projective special linear group = $SL(2,q)/\langle -I\rangle$. $\lvert PSL(2,q)\rvert = \lvert SL(2,q)\rvert/2$.

To simplify expressions in later sections, let $\mathcal{K}_n \triangleq \{0,1,\ldots,n-2\}$ for integers $n \geq 2$.

III. Ingleton Violation: Computer Search and Some Conditions

Since the Ingleton inequality (4) involves four subgroups of a finite group and their various intersections, designing a small admissible structure is very difficult without an existing example. So we use computer programs to search for a small instance. Specifically, we use the GAP system [27] to search its "Small Group" library, which contains all finite groups of order less than or equal to 2000, except those of 1024. We pick a group in this library (starting from the smallest, of course), find all its subgroups, then test the Ingleton inequality for all 4-combinations of these subgroups. This is a tremendous task, as there are already more than 1000 groups (up to isomorphism) of order less than or equal to 100, each of which might have hundreds of subgroups, or even more.

It is therefore extremely critical to prune our search. In fact, we used the following conditions to exclude groups or subgroups in the search, each of which guarantees that Ingleton is satisfied.

Condition 1: G is abelian. [19]

Condition 2: $G_i \trianglelefteq G$, $\forall i$. [28]

Condition 3: $G_1G_2 = G_2G_1$, or equivalently $G_1G_2 \leq G$.

Condition 4: $G_i = 1$ or G, for some i.

Condition 5: $G_i = G_j$ for some distinct i and j.

Condition 6: $G_{12} = 1$.

Condition 7: $G_i \leq G_j$ for some distinct i and j.

Note that Condition 2 subsumes Condition 1, while Condition 3 subsumes Condition 2. Also Conditions 4 and 5 are contained in Condition 7. Nevertheless, we still list these more restrictive conditions as they are easier to check using computer programs. In addition, Conditions 1, 3 and 6 are crucial in our program, as they appear in the outer loops and can save a large amount of search work.

For the above reasons we only list the proofs for Conditions 3, 6 and 7 below:

Proof 3: Construct random variables X_i's from uniformly distributed Λ on G as in Section I-A. As $G_{1;2} \triangleq G_1G_2 \leq G$, we can similarly construct random variable $Z = \Lambda G_{1;2}$. In fact Z is a common information for X_1 and X_2: since $|G_{1;2}| = |G_1||G_2|/|G_{12}|$,

$$H(Z) = H(X_1) + H(X_2) - H(X_1, X_2) = I(X_1; X_2).$$

Also $H(Z|X_1) = H(Z|X_2) = 0$ as $G_1, G_2 \leq G_{1;2}$. Thus Ingleton is satisfied by Lemma 1. ■

In the proof above we used the group-entropy correspondence in Section I-A to translate the problem to the entropy domain. Henceforth, in order to show that a group satisfies Ingleton, we shall either prove (4) directly, or equivalently prove (3) using this correspondence. Furthermore, observe that the Ingleton inequality has symmetries between subscripts 1 and 2 and between 3 and 4, i.e. if we interchange the subscripts 1 and 2, or 3 and 4, the inequality stays the same. Thus if we prove conditions for some $i \in \{1, 2\}$ and $j \in \{3, 4\}$, we automatically get conditions for all $(i, j) \in \{1, 2\} \times \{3, 4\}$. So without loss of generality, we will just prove conditions for the case $(i, j) = (1, 3)$ when these symmetries apply.

Proof 6: Realize that (3) can be rewritten as

$$\delta_{13,14} + \delta_{23,24} + \delta_{134,234} - \delta_{123,124} \geq 0, \tag{5}$$

where for $\emptyset \neq \alpha, \beta \subseteq \mathcal{N}$,

$$\delta_{\alpha,\beta} \triangleq h_\alpha + h_\beta - h_{\alpha\cap\beta} - h_{\alpha\cup\beta}.$$

For example, $\delta_{134,234} = h_{134} + h_{234} - h_{34} - h_{1234}$. By submodularity of entropies, all $\delta_{\alpha,\beta} \geq 0$. If $G_{12} = 1$, then $\delta_{123,124} = 0$ and (5) holds. ■

Proof 7: $(i,j)=(1,2)$ implies Condition 3. $(i,j)=(1,3)$ implies $\delta_{123,124}=0$ in (5). $(i,j)=(3,1)$ implies $\delta_{123,234}=0$ and so $\delta_{123,234}\leq\delta_{12,24}$, which further transforms to $\delta_{123,124}\leq\delta_{23,24}$, thus (5) holds. For $(i,j)=(3,4)$, (4) becomes

$$|G_1||G_2||G_3||G_{123}||G_{124}|\geq|G_{12}||G_{13}||G_{14}||G_{23}||G_{24}|,$$

which is true as $G_2\geq G_{24}$ and by submodularity, $|G_1||G_{124}|\geq|G_{12}||G_{14}|$ and $|G_3||G_{123}|\geq|G_{13}||G_{23}|$. ■

IV. The Smallest Violation Instance and the Group Presentation

Using GAP we found the smallest group that violates Ingleton is $G=S_5$, which has 60 sets of violating subgroups up to subscript symmetries. Further examination shows that these 60 sets of subgroups are in fact all conjugates of each other, thus are virtually the same in terms of group structure. We list below some information from GAP about one representative:[3]

$$\begin{aligned}
&G_1=\big\langle(3,4,5),(1,2)(4,5)\big\rangle && \cong S_3\cong D_6 && |G_1|=6\\
&G_2=\big\langle(1,2,3,4,5),(1,4,3,5)\big\rangle && \cong\mathbb{Z}_5\rtimes\mathbb{Z}_4 && |G_2|=20\\
&G_3=\big\langle(2,3),(1,3,4,2)\big\rangle && \cong D_8 && |G_3|=8\\
&G_4=\big\langle(2,4),(1,2,5,4)\big\rangle && \cong D_8 && |G_4|=8\\
&G_{12}=\big\langle(1,2)(3,5)\big\rangle && \cong\mathbb{Z}_2 && |G_{12}|=2\\
&G_{13}=\big\langle(1,2)(3,4)\big\rangle && \cong\mathbb{Z}_2 && |G_{13}|=2\\
&G_{14}=\big\langle(1,2)(4,5)\big\rangle && \cong\mathbb{Z}_2 && |G_{14}|=2\\
&G_{23}=\big\langle(1,3,4,2)\big\rangle && \cong\mathbb{Z}_4 && |G_{23}|=4\\
&G_{24}=\big\langle(1,2,5,4)\big\rangle && \cong\mathbb{Z}_4 && |G_{24}|=4\\
&G_{34}=1 && && |G_{34}|=1\\
&G_{123}=1 && && |G_{123}|=1\\
&G_{124}=1 && && |G_{124}|=1.
\end{aligned}$$

Simple calculation shows that

$$|G_1||G_2||G_{34}||G_{123}||G_{124}|=120<128=|G_{12}||G_{13}||G_{14}||G_{23}||G_{24}|,$$

so Ingleton is violated. Also we can check that G_1–G_4 indeed generate G.

[3]The permutations are written in cycle notation, e.g. $(1,2)(3,4,5)$ is the permutation on the set $\{1,2,3,4,5\}$ that makes the following mapping: $1\mapsto 2,\ 2\mapsto 1,\ 3\mapsto 4,\ 4\mapsto 5,\ 5\mapsto 3$. Also GAP's convention for permutations is used throughout this paper, i.e. permutations are applied to an element from the right.

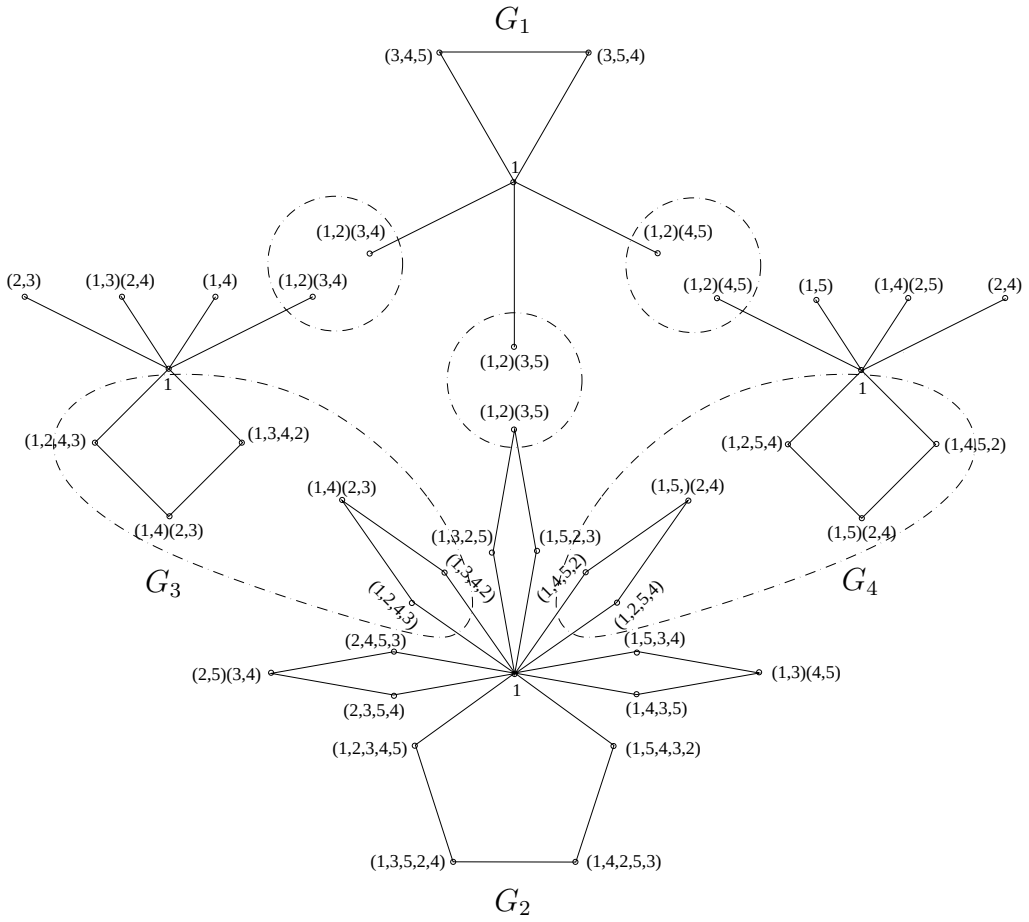


Fig. 1. The cycle graph of the Ingleton violating subgroups of S_5

To illustrate the structure of these subgroups, we use the group cycle graph. See Fig. 1, where the dash-dotted lines denote the pairwise intersections of subgroups excluding identity. From the cycle graph we can obtain more structural information which GAP does not show us directly. First, not only is G_2 a semidirect product of two cyclic groups $\langle(1,2,3,4,5)\rangle \cong \mathbb{Z}_5$ and $\langle(1,4,3,5)\rangle \cong \mathbb{Z}_4$, but also $\big(G_2 \setminus \langle(1,2,3,4,5)\rangle\big) \bigcup \{1\}$ is the union of subgroups which are all isomorphic to (in fact, conjugate to) $\langle(1,4,3,5)\rangle$ and have trivial pairwise intersections. We say G_2 has a "flower" structure in this case. Second, G_4 is the conjugate of G_3 by $(3,4,5)$. In particular, there is a conjugacy relation between the order-4 generators of G_3 and G_4: $(1,3,4,2)^{(3,4,5)} = (1,4,5,2) = (1,2,5,4)^{-1}$.

In order to generalize these subgroups to a family of violations, we seek a parameterized group presentation for G that retains the above structures. Although these group presentations are abstract, each of them can be input to GAP to yield an isomorphic concrete group, and Ingleton inequality can be

checked against the corresponding subgroups. Observing that $|G_{23}|$ and $|G_{24}|$ (both equal to 4) contribute most to the right-hand side (RHS) of (4), we may try to let the "petals" of G_2 (conjugates of $\langle (1,4,3,5) \rangle$) grow while keeping other structures fixed.[4] In the rest of this section, we start from a presentation of G_2 and then extend it to the whole group G.

Let us first define a presentation of a group. For a precise definition one needs to introduce the concept of free groups, which we will skip. The interested readers may consult abstract algebra textbooks, e.g. [26], [29]. Here we only give an informal but useful definition.

Definition 1 (Group Presentation): A set S of *generators* of a group G is a subset of G, such that every group element can be written as a finite product of elements of S and their inverses. An equation satisfied in G involving only $S \cup \{1\}$ is called a *relation* in G among S. Let R be a set of such relations. We say G has a *presentation*

$$\langle\, S \mid R \,\rangle$$

if G is the largest ("freest") group generated by S subject only to the relations R. (Formally, the group G is said to have the above presentation if it is isomorphic to the quotient of a free group F on S by the normal subgroup of F generated by the relations R.)

For example, consider a presentation $\langle x \mid x^n = 1 \rangle$. Any group generated by x contains only the powers of x, but by the relation $x^n = 1$ the order of such a group cannot exceed n. Among these groups the cyclic group $\mathbb{Z}_n$ has the maximum order, hence has the above group presentation.

A. Presentation of G_2

Let G_2 be generated by two elements a and b, with a normal subgroup $N = \langle a \rangle \cong \mathbb{Z}_n$ and another subgroup $H = \langle b \rangle \cong \mathbb{Z}_m$, for some integers m, n. This gives us a presentation

$$G_2 = \left\langle a, b \,\middle|\, a^n = b^m = 1,\, a^b = a^s \right\rangle \tag{6}$$

for some $0 < s < n$. In order to violate Ingleton as much as possible, we may wish for n to be small while m is large. However, the flower structure of G_2 may limit the choices of n and m. First of all, for this presentation to be a semidirect product, we need $s^m \equiv 1 \pmod{n}$ (see [29, Sec 5.4]), i.e.,

$$s \in \mathbb{Z}_n^\times, \quad |s| \,\middle|\, m, \tag{7}$$

[4]This approach is a little conservative, but it is the only successful extension according to our GAP trials. For example, one may try to expand G_1 at the same time, but the structures of G_3 and G_4 usually collapse.

where $|s|$ denotes the order of s in the multiplicative group $\mathbb{Z}_n^\times$. As a consequence, $|G_2| = mn$, $H \bigcap N = 1$, and by the relations in (6) we also have

$$(a^i)^{b^k} = a^{is^k}, \quad \forall i, k \in \mathbb{Z}. \tag{8}$$

Moreover, we need $(G_2 \setminus N) \bigcup \{1\}$ to be the union of subgroups which are all isomorphic to H with trivial pairwise intersections.

One possible way to achieve this is to restrict $H^{g_1} \bigcap H^{g_2} = 1,\ \forall g_1 \neq g_2 \in N$, as in our original example. This is equivalent to $H^g \bigcap H = 1,\ \forall g \in N \setminus \{1\}$. If this is the case, then there will be $|N| = n$ "petals" of size m in G_2, and the total number of nonidentity elements will equal $n(m-1) = nm - n = |G_2 \setminus N|$, and then indeed the flower structure would be achieved. Pick two nonidentity elements $h_1 = b^l \in H$, $h_2 = (b^k)^{a^i} \in H^{a^i}$ for some $0 < k, l < m$ and some $0 < i < n$. Then

$$h_1 = h_2 \ \Leftrightarrow\ a^{-i}b^k a^i = b^l \ \Leftrightarrow\ a^{-i}(a^i)^{b^{-k}}b^k = b^l \ \Leftrightarrow\ a^{-i}a^{is^{-k}} = b^{l-k} \ \Leftrightarrow\ a^{(s^{-k}-1)i} = b^{l-k}.$$

In the last equation, $LHS \in N$ and $RHS \in H$. But $H \bigcap N = 1$ forces that $a^{(s^{-k}-1)i} = b^{l-k} = 1$, i.e. $l = k$ and $n \,\big|\, (s^{-k}-1)i$.

To guarantee that $H^{a^i} \bigcap H = 1$, we must have $m \leq |s|$. Otherwise if we let $0 < k = |s| < m$, then $s^{-k} \equiv 1 \pmod{n}$ and so $n \,\big|\, (s^{-k}-1)i$ is satisfied. This means that by choosing $k = l = |s|$, we have found a nonidentity element $h_2 = (b^k)^{a^i} = b^l = h_1$ in $H^{a^i} \bigcap H$. Therefore $m \leq |s|$ and as $|s| \,\big|\, m$ by (7), $m = |s|$. In particular, $m \leq |\mathbb{Z}_n^\times| \leq n-1$.

For m to be as large as possible, s should be a primitive root modulo n, which makes $m = |\mathbb{Z}_n^\times|$. Pick $n = p$ for some prime p, then $m = |\mathbb{Z}_p^\times| = p-1$ achieves the upper bound $m \leq n-1$. Also in this case, if $0 < k < m = |s|$ and $0 < i < n = p$, then $n \,\big|\, (s^{-k}-1)i$ requires $p \,\big|\, i$ or $p \,\big|\, (s^{-k}-1)$. Since $p > i$, the latter must be true, which implies that $|s| \,\big|\, k$. But this is a contradiction since $0 < k < |s|$. So indeed we have $H^g \bigcap H = 1$, $\forall g \in N$, and the flower structure is realized. Furthermore, to make H nontrivial, we need $p > 2$. Thus with such a choice of parameters, the presentation of G_2 becomes

$$G_2 = \left\langle a, b \,\middle|\, a^p = b^{p-1} = 1,\ a^b = a^s \right\rangle, \tag{9}$$

where $p > 2$ is a prime and s is a primitive root modulo p.

B. Presentation of G

The next step is to extend the presentation (9) to the whole group G generated by G_1–G_4, with the structure in Fig. 1. Consider the dihedral groups G_3 and G_4. The subgroups of rotations are just H^{a_3} and H^{a_4}, respectively, for some $a_3 = a^{k_3}, a_4 = a^{k_4} \in N$. Also G_3 and G_4 each shares one element of

TABLE I

CORRESPONDENCE OF GROUP ELEMENTS

a	b	c	b_1	b_3	b_4
$(1,2,3,4,5)$	$(1,4,3,5)$	$(3,4,5)$	$(1,2)(3,5)$	$(1,3,4,2)$	$(1,2,5,4)$

reflection with the dihedral group G_1, while the remaining reflection of G_1 is just $(b^{\frac{p-1}{2}})^{a_1}$ in G_2, for some $a_1 = a^{k_1} \in N$. Thus if we can determine the generator of the subgroup of rotations of G_1, then all elements of G_1–G_4 are determined. In other words, if we introduce an element c as the generator of rotations of G_1, then all elements from G_1–G_4 can be express as products of a, b, c and their inverses. Define

$$b_1 = (b^{\frac{p-1}{2}})^{a^{k_1}}, \quad b_3 = b^{a^{k_3}}, \quad b_4 = b^{a^{k_4}} \tag{10}$$

for some integers k_1, k_3, k_4. If in Fig. 1 we let a, b, c, b_1, b_3, b_4 correspond with the elements specified in Table I, then the subgroups and the whole group in our presentation should be

$$G_1 = \langle c, b_1 \rangle, \quad G_2 = \langle a, b \rangle, \quad G_3 = \langle b_1 c^2, b_3 \rangle, \quad G_4 = \langle b_1 c, b_4 \rangle, \quad G = \langle a, b, c \rangle. \tag{11}$$

As $G_1 \cong D_6$, we should have the relation $c^3 = (cb_1)^2 = 1$. Furthermore, for G_3 and G_4 to be dihedral groups, we need $(b_3 \cdot b_1 c^2)^2 = (b_4 \cdot b_1 c)^2 = 1$.

At this point we can try to plug in the presentation with these relations to GAP to find a concrete group. But still there are too many parameters to choose, especially when p is large, the choices of k_1, k_3, k_4 are numerous. Also for a fixed p not many such combinations yield successful Ingleton violations, according to our GAP trials. Therefore we need to ultilize more structural information from Fig. 1 to obtain more restrictions on k_1, k_3 and k_4.

Observe that in the original violation, G_4 is the conjugate of G_3 by $(3,4,5)$, and $(1,3,4,2)^{(3,4,5)} = (1,2,5,4)^{-1}$. In our presentation this translates to $b_3^c = b_4^{-1}$, according to Table I. With this new relation, we claim that $(b_3 \cdot b_1 c^2)^2 = (b_4 \cdot b_1 c)^2 = 1$ is satisfied if and only if $k_3 - k_1 \equiv k_1 - k_4 \pmod{p}$. In fact, as $|b_1| = 2$, $c^3 = (cb_1)^2 = 1$, we have $cb_1 = b_1 c^2$ and $b_1 c = c^2 b_1$. Using these relations we can establish the following equalities:

$$(b_3 \cdot b_1 c^2)^2 = b_3 b_1 c^{-1} b_3 c b_1 = b_3 b_1 b_4^{-1} b_1,$$

$$(b_4 \cdot b_1 c)^2 = b_4 b_1 c b_4 c^{-1} b_1 = b_4 b_1 b_3^{-1} b_1 = \left((b_3 b_1 b_4^{-1} b_1)^{-1}\right)^{b_1}.$$

So $(b_3 \cdot b_1 c^2)^2 = 1$ if and only if $(b_4 \cdot b_1 c)^2 = 1$. Using (8) and the fact that $b^{\frac{p-1}{2}} = (b^{\frac{p-1}{2}})^{-1}$ and plugging (10) in, we have

$$\begin{aligned}
b_3 b_1 b_4^{-1} b_1 &= b^{a^{k_3}} (b^{\frac{p-1}{2}})^{a^{k_1}} (b^{-1})^{a^{k_4}} (b^{\frac{p-1}{2}})^{a^{k_1}} \\
&= a^{-k_3} b a^{k_3-k_1} b^{\frac{p-1}{2}} a^{k_1-k_4} b^{-1} a^{k_4-k_1} b^{\frac{p-1}{2}} a^{k_1} \\
&= a^{-k_3} \cdot b a^{k_3-k_1} b^{-1} \cdot b^{\frac{p-1}{2}} \cdot b a^{k_1-k_4} b^{-1} \cdot a^{k_4-k_1} b^{\frac{p-1}{2}} a^{k_1} \\
&= a^{-k_3} \cdot a^{(k_3-k_1)s^{-1}} \cdot b^{\frac{p-1}{2}} \cdot a^{(k_1-k_4)s^{-1}} \cdot a^{k_4-k_1} b^{\frac{p-1}{2}} a^{k_1} \\
&= a^{(k_3-k_1)s^{-1}-k_3} \cdot (b^{\frac{p-1}{2}})^{-1} a^{(k_1-k_4)(s^{-1}-1)} b^{\frac{p-1}{2}} \cdot a^{k_1} \\
&= a^{(k_3-k_1)s^{-1}-k_3} \cdot a^{(k_1-k_4)(s^{-1}-1)s^{(p-1)/2}} \cdot a^{k_1} \\
&= a^{[(k_3-k_1)+(k_1-k_4)s^{(p-1)/2}](s^{-1}-1)}.
\end{aligned}$$

Since s is a primitive root modulo p, $|s^{(p-1)/2}| = 2$. As $\mathbb{Z}_p^\times$ is cyclic of an even order $p-1$, it is clear that there is a unique element of order 2. But -1 has order 2 in $\mathbb{Z}_p^\times$, so $s^{(p-1)/2} \equiv -1 \pmod{p}$ and

$$(b_3 \cdot b_1 c^2)^2 = b_3 b_1 b_4^{-1} b_1 = a^{[(k_3-k_1)-(k_1-k_4)](s^{-1}-1)}.$$

Now $p \nmid (s^{-1}-1)$ as $s \neq 1$, which implies

$$(b_3 \cdot b_1 c^2)^2 = 1 \;\Leftrightarrow\; p \,\big|\, [(k_3-k_1)-(k_1-k_4)] \;\Leftrightarrow\; k_3 - k_1 \equiv k_1 - k_4 \pmod{p}.$$

This condition on k_1, k_3 and k_4 tells us that the petals G_{23} and G_{24} of G_2 should be symmetric (modulo p) w.r.t. G_{12}, i.e. G_{23}, G_{12} and G_{24} should be equally spaced.[5]

In sum, our analysis leads to the following presentation:

$$G = \left\langle a, b, c \,\middle|\, a^p = b^{p-1} = c^3 = 1,\ a^b = a^s,\ (cb_1)^2 = b_3^c b_4 = 1 \right\rangle \tag{12}$$

where p is an odd prime, s is a primitive root modulo p, $k_3 - k_1 \equiv k_1 - k_4 \pmod{p}$. If our extension of the subgroup structures succeeds, then the orders of subgroups and intersections would be: $|G_1| = 6$, $|G_2| = p(p-1)$, $|G_3| = |G_4| = 2(p-1)$, $|G_{12}| = |G_{13}| = |G_{14}| = 2$, $|G_{23}| = |G_{24}| = p-1$, $|G_{34}| = |G_{123}| = |G_{124}| = 1$. Hence LHS of (4) $= 6p(p-1)$ while $RHS = 8(p-1)^2$, and so when $p \geq 5$, Ingleton should be violated.

[5]With this symmetry it is very easy for GAP to produce the desired structures, even with arbitrary choices of k_1 and k_3.

V. Explicit Violation Construction with $PGL(2,p)$ and $PGL(2,q)$

Feeding the above presentation to GAP, we find that for $p = 5, 7, \ldots, 23$ the outcome is a finite group that violates the Ingleton inequality.[6] Moreover, with GAP we verified for the first few primes (up to $p = 11$) that this group is isomorphic to the projective general linear group $PGL(2,p)$. This leads us to conjecturing that $PGL(2,p)$ is a family of Ingleton-violating groups. In fact, with an explicit identification of the generators in (12) with matrices in $PGL(2,p)$, we prove that $PGL(2,p)$ is indeed a family of Ingleton-violating groups for primes $p \geq 5$, by directly constructing their violating subgroups in (11) in the form of matrices. These matrix subgroups all have clear interpretations. Furthermore, once we have the formats of these subroups, we extend them to the Ingleton-violating family $PGL(2,q)$ for all finite field order $q \geq 5$.

A. The Family $PGL(2,p)$

First we introduce some necessary notations. Let p be an odd prime. For $A \in GL(2,p)$, let $\overline{A}$ denote the left coset of A in $GL(2,p)$ with respect to the center $V_p = \{\alpha I : \alpha \in \mathbb{F}_p^\times\}$. Thus $\overline{A} = \overline{B}$ if and only if each entry of A is a nonzero constant multiple of the corresponding entry of B. A^T denotes the transpose of A as usual. We denote the elements of $\mathbb{F}_p$ by ordinary integers, but the addition and multiplication, as well as equality, are modulo p. Furthermore, $-k$ and k^{-1} denotes the additive and multiplicative inverses of k in $\mathbb{F}_p$ respectively. If $s \in \mathbb{F}_p$, and A has multiplicative order p, then A^s simply indicates the s-th power of A, where s is viewed as an integer.

We start by identifying the generators in $PGL(2,p)$ that correspond to presentation (12). Consider the following matrices in $GL(2,p)$:

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 0 \\ 0 & t \end{bmatrix}, \qquad C = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}$$

where t is a primitive root modulo p, i.e. a generator of $\mathbb{F}_p^\times$. Our guess is that $\overline{A}, \overline{B}, \overline{C}$ correspond to the generators a, b, c in (12) respectively. The powers of these matrices are:

$$A^k = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}, \quad B^k = \begin{bmatrix} 1 & 0 \\ 0 & t^k \end{bmatrix}, \quad C^2 = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}, \quad C^3 = I$$

for any integer k. Thus $\left|\overline{A}\right| = p$, $\left|\overline{B}\right| = p-1$, and $\left|\overline{C}\right| = 3$. Also,

$$A^B = B^{-1}AB = \begin{bmatrix} 1 & 0 \\ t^{-1} & 1 \end{bmatrix} = A^s,$$

[6]The capability of the testing program is primarily limited by hardware. When p is too large the program runs out of memory.

where $s = t^{-1}$ is also a primitive root modulo p. So $\overline{A}^{\overline{B}} = \overline{A}^s$. Next we let

$$B_1 = (B^{\frac{p-1}{2}})^{A^{k_1}} = \begin{bmatrix} 1 & 0 \\ -k_1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ k_1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -2k_1 & -1 \end{bmatrix},$$

where we calculated $t^{\frac{p-1}{2}} = -1$ as it is the unique element of order 2 in $\mathbb{F}_p^\times$. Now check

$$CB_1 = \begin{bmatrix} -2k_1 & -1 \\ 2k_1 - 1 & 1 \end{bmatrix}, \quad (CB_1)^2 = \begin{bmatrix} 4k_1^2 - 2k_1 + 1 & 2k_1 - 1 \\ -(2k_1 - 1)^2 & 2 - 2k_1 \end{bmatrix}.$$

Thus if we want $\left(\overline{CB_1}\right)^2 = \overline{I}$, k_1 must be $2^{-1} = \frac{p+1}{2}$. In this case,

$$B_1 = \begin{bmatrix} 1 & 0 \\ -1 & -1 \end{bmatrix}, \quad CB_1 = \begin{bmatrix} -1 & -1 \\ 0 & 1 \end{bmatrix}, \quad \left(\overline{CB_1}\right)^2 = \overline{I}.$$

Let $B_3 = B^{A^{k_3}}$, $B_4 = B^{A^{k_4}}$. As $k_3 - k_1 = k_1 - k_4$, we have $k_3 = 1 - k_4$.

$$B^{A^k} = \begin{bmatrix} 1 & 0 \\ k(t-1) & t \end{bmatrix}, \quad B_3 C \cdot B_4 = \begin{bmatrix} 0 & 1 \\ -t & k_3(t-1) - t \end{bmatrix} \begin{bmatrix} 1 & 0 \\ k_4(t-1) & t \end{bmatrix},$$

whose $(1,1)$-entry is $k_4(t-1)$. If we want $\overline{B_3}^{\overline{C}}\overline{B_4} = \overline{I}$, i.e., $\overline{B_3 C B_4} = \overline{C}$, k_4 must be 0 since the $(1,1)$-entry of C is 0 and $t \neq 1$. So $k_3 = 1 - k_4 = 1$,

$$B_3 = \begin{bmatrix} 1 & 0 \\ t-1 & t \end{bmatrix}, \quad B_4 = \begin{bmatrix} 1 & 0 \\ 0 & t \end{bmatrix} = B, \quad \overline{B_3 C B_4} = \overline{\begin{bmatrix} 0 & 1 \\ -t & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & t \end{bmatrix}} = \overline{\begin{bmatrix} 0 & t \\ -t & -t \end{bmatrix}} = \overline{C}.$$

So far for $\overline{A}, \overline{B}, \overline{C}$ we have verified all the relations in (12). We can also prove that they are actually a set of generators for $PGL(2,p)$. Observe that each matrix in $GL(2,p)$ can be written as a product of some elementary matrices, which are

$$\begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & \beta \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & t^i \end{bmatrix}, \quad \begin{bmatrix} t^j & 0 \\ 0 & 1 \end{bmatrix}$$

where $\alpha, \beta \in \mathbb{F}_p$ and $i, j \in \mathcal{K}_p$. They are generated by A, A^T, B and $t^{-1}B$ respectively. So $PGL(2,p)$ is generated by $\overline{A}, \overline{A^T}$ and $\overline{B}$. Now check

$$B_1 C = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad A^{B_1 C} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = A^T,$$

thus $\overline{A}, \overline{B}$ and $\overline{C}$ generate $PGL(2,p)$. Hence setting $s = t^{-1}$, $k_1 = \frac{p+1}{2}$, $k_3 = 1$, $k_4 = 0$, we see that $PGL(2,p)$ is a quotient of the group G in (12), whose generators $\overline{A}, \overline{B}$ and $\overline{C}$ correspond precisely to the generators a, b and c of G.

Remark 1: Note that we have not proved that (12) is a presentation of $PGL(2,p)$. To do that, one must show that the order of the group generated by a, b, c in (12) is no more than $|PGL(2,p)| = (p-1)p(p+1)$, which we have not yet been able to prove. However, identifying possible corresponding generators still gives us a way to explicitly construct the subgroups to violate Ingleton.

Now we can write out the subgroups in $PGL(2,p)$ corresponding to subgroups in (11).

$G_1 = \langle \overline{C}, \overline{B_1} \rangle$. Note that $\left|\overline{C}\right| = 3$, $\left|\overline{B_1}\right| = 2$, and $\left(\overline{CB_1}\right)^2 = \overline{I}$, so $\overline{CB_1} = \overline{B_1}\left(\overline{C}\right)^2$ and G_1 has at most 6 elements $\{\left(\overline{B_1}\right)^i \left(\overline{C}\right)^j : 0 \le i < 2,\ 0 \le j < 3\}$. Calculating these elements we can see $|G_1| = 6$ exactly and thus indeed $G_1 \cong D_6 \cong S_3$:

$$G_1 = \left\{ \overline{I},\ \overline{\begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}},\ \overline{\begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}},\ \overline{\begin{bmatrix} 1 & 0 \\ -1 & -1 \end{bmatrix}},\ \overline{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}},\ \overline{\begin{bmatrix} -1 & -1 \\ 0 & 1 \end{bmatrix}} \right\}.$$

$G_2 = \langle \overline{A}, \overline{B} \rangle$. We claim that G_2 is the subgroup of lower triangular matrices[7] in $GL(2,p)$ modulo V_p, i.e.,

$$G_2 = \left\{ \overline{\begin{bmatrix} 1 & 0 \\ \alpha & \beta \end{bmatrix}} \;\middle|\; \alpha \in \mathbb{F}_p,\ \beta \in \mathbb{F}_p^{\times} \right\}.$$

As A, B are lower triangular, any element in G_2 is a lower triangular matrix modulo V_p. On the other hand, $\forall \alpha \in \mathbb{F}_p,\ \beta \in \mathbb{F}_p^{\times}$, then $\beta = t^l$ for some integer l. So

$$\begin{bmatrix} 1 & 0 \\ \alpha & \beta \end{bmatrix} = A^{\alpha} B^{l} \Rightarrow \overline{\begin{bmatrix} 1 & 0 \\ \alpha & \beta \end{bmatrix}} = \overline{A}^{\alpha} \overline{B}^{l} \in G_2.$$

Thus $|G_2| = p(p-1)$ and G_2 has presentation (9). Therefore, as proved in Section IV-A, $G_2 \cong \mathbb{Z}_p \rtimes \mathbb{Z}_{p-1}$ and it achieves the desired flower structure.

$G_3 = \langle \overline{B_1}\left(\overline{C}\right)^2, \overline{B_3} \rangle = \langle \overline{CB_1}, \overline{B_3} \rangle$. Note that $\left|\overline{CB_1}\right| = 2$, $\left|\overline{B_3}\right| = \left|\overline{B}\right| = p-1$, also

$$B_3^k = \begin{bmatrix} 1 & 0 \\ t^k - 1 & t^k \end{bmatrix}, \quad B_3^{-1} = \begin{bmatrix} 1 & 0 \\ t^{-1} - 1 & t^{-1} \end{bmatrix},$$

$$\overline{B_3} \cdot \overline{CB_1} = \overline{\begin{bmatrix} -1 & -1 \\ 1 - t & 1 \end{bmatrix}} = \overline{\begin{bmatrix} -t^{-1} & -t^{-1} \\ t^{-1} - 1 & t^{-1} \end{bmatrix}} = \overline{CB_1}\left(\overline{B_3}\right)^{-1},$$

[7]We would end up with upper triangular matrices for G_2 if A^T were used in place of A. But the two resulting groups are actually conjugate to each other, e.g. consider conjugating by $B_1 C$.

so G_3 has at most $2(p-1)$ elements $\{\left(\overline{CB_1}\right)^i\left(\overline{B_3}\right)^j : 0 \le i < 2,\ 0 \le j < p-1\}$. Calculating these elements we can see $|G_3| = 2(p-1)$ exactly and so $G_3 \cong D_{2(p-1)}$:

$$G_3 = \left\{ \left(\overline{B_3}\right)^k = \overline{\begin{bmatrix} 1 & 0 \\ t^k - 1 & t^k \end{bmatrix}},\quad \overline{CB_1}\left(\overline{B_3}\right)^k = \overline{\begin{bmatrix} -1 & -1 \\ 1 - t^{-k} & 1 \end{bmatrix}} \;\middle|\; k \in \mathcal{K}_p \right\}.$$

$G_4 = \left\langle \overline{B_1C}, \overline{B_4} \right\rangle$. Note that $\left|\overline{B_1C}\right| = 2$, $\left|\overline{B_4}\right| = \left|\overline{B}\right| = p-1$. Moreover,

$$\overline{B_4} \cdot \overline{B_1C} = \overline{\begin{bmatrix} 0 & 1 \\ t & 0 \end{bmatrix}} = \overline{\begin{bmatrix} 0 & t^{-1} \\ 1 & 0 \end{bmatrix}} = \overline{B_1C}\left(\overline{B_4}\right)^{-1},$$

so G_4 has at most $2(p-1)$ elements $\{\left(\overline{B_1C}\right)^i\left(\overline{B_4}\right)^j : 0 \le i < 2,\ 0 \le j < p-1\}$. Calculating these elements we can see $|G_4| = 2(p-1)$ exactly and so $G_4 \cong D_{2(p-1)}$:

$$G_4 = \left\{ \left(\overline{B_4}\right)^k = \overline{\begin{bmatrix} 1 & 0 \\ 0 & t^k \end{bmatrix}},\quad \overline{B_1C}\left(\overline{B_4}\right)^k = \overline{\begin{bmatrix} 0 & t^k \\ 1 & 0 \end{bmatrix}} \;\middle|\; k \in \mathcal{K}_p \right\}.$$

These are all diagonal and anti-diagonal matrices in $GL(2,p)$ modulo V_p. Note that we have already verified $\left(\overline{B_3}\right)^{\overline{C}} = \overline{B_4}^{-1}$, also $\left(\overline{CB_1}\right)^{\overline{C}} = \overline{B_1C}$, thus indeed $G_4 = G_3^{\overline{C}}$ as in the original instance (Fig. 1).

With all four subgroups explicitly written, we can easily write down the intersections:

$$G_{12} = \left\langle \overline{B_1} \right\rangle = \left\{ \overline{I},\quad \overline{\begin{bmatrix} 1 & 0 \\ -1 & -1 \end{bmatrix}} \right\} \cong \mathbb{Z}_2,\quad G_{13} = \left\langle \overline{CB_1} \right\rangle = \left\{ \overline{I},\quad \overline{\begin{bmatrix} -1 & -1 \\ 0 & 1 \end{bmatrix}} \right\} \cong \mathbb{Z}_2,$$

$$G_{14} = \left\langle \overline{B_1C} \right\rangle = \left\{ \overline{I},\quad \overline{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}} \right\} \cong \mathbb{Z}_2,\quad G_{23} = \left\langle \overline{B_3} \right\rangle = \left\{ \overline{\begin{bmatrix} 1 & 0 \\ t^k - 1 & t^k \end{bmatrix}} \;\middle|\; k \in \mathcal{K}_p \right\} \cong \mathbb{Z}_{p-1},$$

$$G_{24} = \left\langle \overline{B_4} \right\rangle = \left\{ \overline{\begin{bmatrix} 1 & 0 \\ 0 & t^k \end{bmatrix}} \;\middle|\; k \in \mathcal{K}_p \right\} \cong \mathbb{Z}_{p-1},\quad G_{34} = G_{123} = G_{124} = 1.$$

$$|G_{12}| = |G_{13}| = |G_{14}| = 2,\quad |G_{23}| = |G_{24}| = p-1.$$

So in (4), indeed $LHS = |G_1||G_2||G_{34}||G_{123}||G_{124}| = 6p(p-1)$ and $RHS = |G_{12}||G_{13}||G_{14}||G_{23}||G_{24}| = 8(p-1)^2$, hence $LHS - RHS = 2(p-1)(4-p)$. Thus Ingleton is violated when $p \ge 5$, and the subgroup structures of $S_5 \cong PGL(2,5)$ are exactly reproduced.

B. The Family $PGL(2,q)$

With the explicit matrix forms of the Ingleton-violating subgroups, we can further extend the above violation to $PGL(2,q)$, for all finite field order $q \geq 5$. For a finite field $\mathbb{F}_q$, we know that $q = p^m$ for some prime p (the characteristic of $\mathbb{F}_q$) and some integer m. Since $\mathbb{F}_p$ is the prime subfield of $\mathbb{F}_q$, $GL(2,p)$ is a subgroup of $GL(2,q)$, which induces an isomorphic copy of $PGL(2,p)$ as a subgroup of $PGL(2,q)$. Therefore, using the same subgroups of $PGL(2,p)$ as in the previous section, we obtain a trivial Ingleton violation in $PGL(2,q)$ whenever the characteristic $p \geq 5$. Nevertheless, by extending the interpretations of these subgroups to $PGL(2,q)$, we can obtain a more general (nontrivial) violation, for each finite field order $q \geq 5$.

In the field $\mathbb{F}_q$, we continue to use the ordinary integers with modular arithmetic to represent the prime subfield $\mathbb{F}_p$. With this convention, all the matrices and subgroups in Section V-A are well defined[8], although now the cosets are taken with respect to V_q rather than V_p. These subgroups constitute a trivial embedding of our previous violation in $PGL(2,q)$. However, in $PGL(2,q)$, the previous sets of generators do not guarantee that G_2 is the full subgroup of all lower triangular matrices, nor that G_4 contains all the diagonal and anti-diagonal matrices.

To preserve these interpretations of the subgroups, we need to make some adjustment to the generators of G_2. Redefine t to be a primitive element of $\mathbb{F}_q$, i.e. t generates $\mathbb{F}_q^\times$. Then $\left|\overline{B}\right| = q-1$. Also instead of a single A, we need to introduce more matrices to generate the subgroup $N \triangleq \left\{ \overline{A_\alpha} \,\middle|\, \alpha \in \mathbb{F}_q \right\}$, where for each $\alpha \in \mathbb{F}_q$ we define

$$A_\alpha = \begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix}.$$

Clearly $A_\alpha A_\beta = A_{\alpha+\beta}$, and $A_\alpha^k = A_{k\alpha}$ for each integer k. Thus $\left|\overline{A_\alpha}\right| = p$ for each $\alpha \in \mathbb{F}_q^\times$. Observe that $\mathbb{F}_q$ is an m-dimensional vector space over $\mathbb{F}_p$, let $(\xi_1, \xi_2, \ldots, \xi_m)$ be a basis. Then $\forall \alpha \in \mathbb{F}_q$, $\alpha = \sum_{i=1}^m k_i \xi_i$ for some $k_1, k_2, \ldots, k_m \in \mathbb{F}_p$ and $A_\alpha = \prod_{i=1}^m A_{\xi_i}^{k_i}$. Also $\left\langle \overline{A_{\xi_i}} \right\rangle \bigcap \left\langle \overline{A_{\xi_j}} \right\rangle = 1$ for distinct i and j. Thus

$$N = \left\langle\ \overline{A_{\xi_1}},\ \overline{A_{\xi_2}},\ \ldots,\ \overline{A_{\xi_m}}\ \right\rangle \cong \left\langle\ \overline{A_{\xi_1}}\ \right\rangle \times \left\langle\ \overline{A_{\xi_2}}\ \right\rangle \times \ldots \times \left\langle\ \overline{A_{\xi_m}}\ \right\rangle \cong \mathbb{Z}_p^m.$$

Actually, N is isomorphic to the additive group of the vector space $\mathbb{F}_q$ over $\mathbb{F}_p$ (Also see Section VIII-A).

Let $G_2 = \left\langle \overline{A_{\xi_1}}, \overline{A_{\xi_2}}, \ldots, \overline{A_{\xi_m}}, \overline{B} \right\rangle = \left\langle N, \overline{B} \right\rangle$. Similar to the previous section, it is easy to show that now G_2 is indeed the subgroup of all lower triangular matrices modulo V_q. Furthermore, for any

[8]The only problem that may arise is when $p = 2$, $B_1 = (B^{\frac{p-1}{2}})^{A^{k_1}}$ is not well defined. But we can circumvent that by directly working with the final matrix form of B_1.

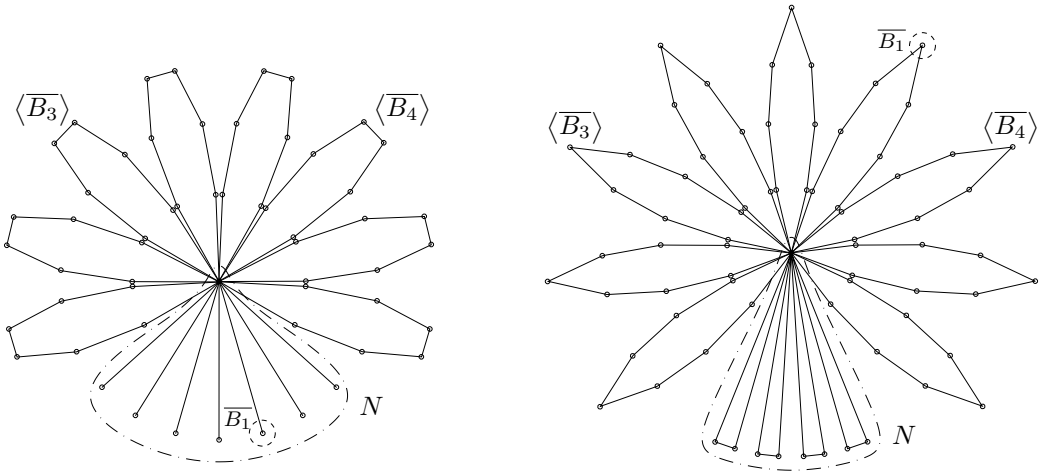


Fig. 2. The generalized flower structures. The center point of each cycle graph denotes the identity element.

$\alpha \in \mathbb{F}_q$, we have $\overline{A_\alpha}^{\overline{B}} = \overline{A_{t^{-1}\alpha}}$, so $N \trianglelefteq G_2$ and $G_2 = NH$, where $H \triangleq \langle\overline{B}\rangle$. Also $N \bigcap H = 1$, thus $G_2 \cong N \rtimes H \cong \mathbb{Z}_p^m \rtimes \mathbb{Z}_{q-1}$. Although in general G_2 does not have presentation (6) or (9) anymore since N is not necessarily cyclic, we can prove that it does have a "generalized flower structure" when $q > 2$, i.e. $(G_2 \setminus N) \bigcup \{\overline{I}\}$ is the union of subgroups which are all isomorphic to H with trivial pairwise intersections. Similar to the analysis of the G_2 in Section IV-A, it suffices to show that $H^{\overline{A_\alpha}} \bigcap H = 1$, $\forall \overline{A_\alpha} \in N \setminus \{\overline{I}\}$. But this is true since for each $\alpha \in \mathbb{F}_q^\times$ and some integers $k, l \in \mathcal{K}_q$,

$$(\overline{B}^k)^{\overline{A_\alpha}} = \overline{B}^l \iff \overline{B}^k \cdot \overline{A_\alpha} = \overline{A_\alpha} \cdot \overline{B}^l \iff \overline{\begin{bmatrix} 1 & 0 \\ t^k\alpha & t^k \end{bmatrix}} = \overline{\begin{bmatrix} 1 & 0 \\ \alpha & t^l \end{bmatrix}} \iff k = l = 0.$$

Fig. 2 shows two representative generalized flower structures of G_2, for $q = 8$ and $q = 9$. In each cycle graph of G_2, there are $|N| = q$ petals and one "root system" (encircled by the dash-dotted line), which is the normal subgroup N. Every petal is a conjugate of H and has size $q-1$. Since N has $q-1$ nonidentity elements, each having order p, the root system consists of $(q-1)/(p-1)$ trivially intersecting "roots/tubers", each of which is a p-cycle. Note that when $m = 1$, there is only one root/tuber, as in the original flower structure in Fig. 1.

Now using the same matrices

$$C = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 1 & 0 \\ -1 & -1 \end{bmatrix}, \quad B_3 = B^{A_1} = \begin{bmatrix} 1 & 0 \\ t-1 & t \end{bmatrix}, \quad B_4 = B = \begin{bmatrix} 1 & 0 \\ 0 & t \end{bmatrix}$$

as in Section V-A (except that t now generates $\mathbb{F}_q^\times$ instead of $\mathbb{F}_p^\times$), we write down the following subgroups:

$G_1 = \langle \overline{C}, \overline{B_1} \rangle \cong D_6 \cong S_3$. (Same as in Section V-A.)

$G_2 = \langle \overline{A_{\xi_1}}, \overline{A_{\xi_2}}, \ldots, \overline{A_{\xi_m}}, \overline{B} \rangle = \langle N, \overline{B} \rangle \cong \mathbb{Z}_p^m \rtimes \mathbb{Z}_{q-1}$, which consists of all lower triangular matrices in $GL(2,q)$ modulo V_q.

$G_3 = \langle \overline{B_1}(\overline{C})^2, \overline{B_3} \rangle = \langle \overline{CB_1}, \overline{B_3} \rangle$. Now $|\overline{B_3}| = q-1$, and we still have $\overline{B_3} \cdot \overline{CB_1} = \overline{CB_1}(\overline{B_3})^{-1}$, so

$$G_3 = \left\{ (\overline{B_3})^k = \overline{\begin{bmatrix} 1 & 0 \\ t^k - 1 & t^k \end{bmatrix}}, \quad \overline{CB_1}(\overline{B_3})^k = \overline{\begin{bmatrix} -1 & -1 \\ 1 - t^{-k} & 1 \end{bmatrix}} \;\middle|\; k \in \mathcal{K}_q \right\} \cong D_{2(q-1)}.$$

$G_4 = \langle \overline{B_1 C}, \overline{B_4} \rangle$. Now $|\overline{B_4}| = q-1$ and $\overline{B_4} \cdot \overline{B_1 C} = \overline{B_1 C}(\overline{B_4})^{-1}$, so

$$G_4 = \left\{ (\overline{B_4})^k = \overline{\begin{bmatrix} 1 & 0 \\ 0 & t^k \end{bmatrix}}, \quad \overline{B_1 C}(\overline{B_4})^k = \overline{\begin{bmatrix} 0 & t^k \\ 1 & 0 \end{bmatrix}} \;\middle|\; k \in \mathcal{K}_q \right\} \cong D_{2(q-1)},$$

which comprises all diagonal and anti-diagonal matrices in $GL(2,q)$ modulo V_q.

Next we find the intersections: $G_{12} = \langle \overline{B_1} \rangle$, $G_{13} = \langle \overline{CB_1} \rangle$, and $G_{14} = \langle \overline{B_1 C} \rangle$, which are all isomorphic to $\mathbb{Z}_2$; $G_{23} = \langle \overline{B_3} \rangle$ and $G_{24} = \langle \overline{B_4} \rangle$, both of which are isomorphic to $\mathbb{Z}_{q-1}$; and $G_{34} = G_{123} = G_{124} = 1$.

The orders of the four subgroups are $|G_1| = 6$, $|G_2| = q(q-1)$, $|G_3| = |G_4| = 2(q-1)$, and for the intersections $|G_{12}| = |G_{13}| = |G_{14}| = 2$, $|G_{23}| = |G_{24}| = q-1$, $|G_{34}| = |G_{123}| = |G_{124}| = 1$. So in (4), $LHS = |G_1||G_2||G_{34}||G_{123}||G_{124}| = 6q(q-1)$, while $RHS = |G_{12}||G_{13}||G_{14}||G_{23}||G_{24}| = 8(q-1)^2$. Thus $LHS - RHS = 2(q-1)(4-q)$ and Ingleton is violated when $q \geq 5$.

Remark 2: Depending on the characteristic p of $\mathbb{F}_q$, the intersection $G_{12} = \langle \overline{B_1} \rangle$ might lie in either the petals or the roots of G_2, as depicted by the dashed circles in Fig. 2. If $p \neq 2$, then q is odd and $\overline{B_1} = (\overline{B}^{\frac{q-1}{2}})^{\overline{A_{k_1}}}$ where $k_1 = 2^{-1} = \frac{p+1}{2}$, so G_{12} is on the petal $H^{\overline{A_{k_1}}}$. Whereas if $p = 2$, then $-1 = 1$ and $\overline{B_1} = \overline{A_1} \in N$, so G_{12} becomes a root. Note that the patterns of the other intersections are not changed for different q's.

Remark 3: We can also show that $\overline{A_{\xi_1}}, \overline{A_{\xi_2}}, \ldots, \overline{A_{\xi_m}}, \overline{B}$ and $\overline{C}$ generate $PGL(2,q)$, using the same argument as in the previous section. The only difference is that the elementary matrices of $GL(2,q)$ are now generated by $A_{\xi_1}, A_{\xi_1}^T, \ldots, A_{\xi_m}, A_{\xi_m}^T, B$ and $t^{-1}B$. But as $A_\alpha^{B_1 C} = A_\alpha^T$, $\forall \alpha \in \mathbb{F}_q$, we see that $PGL(2,q)$ is indeed generated by the desired elements.

In Section VII, we will see that these subgroups have more fundamental interpretations in the framework of group actions and groups of Lie type: each subgroup is the stabilizer for a special set of points in the underlying projective geometry of $PGL(2,q)$.

C. Discussion

To measure "how much" the Ingleton inequality is violated, or how effective a set of subgroups is in terms of violating Ingleton, we need to compare the difference of the two sides of (3) for the corresponding entropy vector, i.e.

$$\Delta_h \triangleq h_1 + h_2 + h_{34} + h_{123} + h_{124} - (h_{12} + h_{13} + h_{14} + h_{23} + h_{24}).$$

Translating to the finite group context it equals $\log \frac{RHS}{LHS}$ of (4). Thus we can make the following definition to measure the extent to which Ingleton is violated.

Definition 2: For a 4-tuple of subgroups $\tau = (G_i : 1 \leq i \leq 4)$, we define the *Ingleton ratio* to be

$$r(\tau) = \frac{|G_{12}||G_{13}||G_{14}||G_{23}||G_{24}|}{|G_1||G_2||G_{34}||G_{123}||G_{124}|}. \tag{13}$$

Clearly $\Delta_h = \log r$ and Ingleton is violated iff $r > 1$. The family $PGL(2, q)$ have the Ingleton ratio

$$r = \frac{4(q-1)}{3q},$$

which approaches $4/3$ when q is large.

However, the Ingleton ratio is not precise enough to characterize the effectiveness of an Ingleton violation instance. Observe that $\overline{\Gamma_n^*}$ is a cone, and in fact, as remarked in [23], adding an entropy vector to itself yields another entropy vector. Thus we can arbitrarily increase the Ingleton ratio by joining copies of a violation instance. For example, if $\tau = (G_i : 1 \leq i \leq 4)$ is such an instance, for each integer N let $G' = \times_{k=1}^{N} G \triangleq G \times \cdots \times G$ be the direct product of N copies of G and define $\tau' = (G'_i : 1 \leq i \leq 4)$ with $G'_i = \times_{k=1}^{N} G_i$ for each i. Then the Ingleton ratio $r(\tau') = [r(\tau)]^N$, which grows unbounded when $N \to \infty$.

Therefore we need to consider the scaled version of Δ_h, to be able to measure the effectiveness of an Ingleton violation. In [30] Dougherty *et al.* use the full joint entropy h_{1234} as a scaling factor to avoid the problem above:

Definition 3: For an entropy vector $h = (h_\alpha : \emptyset \neq \alpha \subseteq \{1, 2, 3, 4\})$, define the *Ingleton score* to be

$$\sigma(h) = -\frac{\Delta_h}{h_{1234}}.$$

In the context of groups, the Ingleton score of a 4-tuple τ of subgroups of G is

$$\sigma(\tau) = \frac{-\log r(\tau)}{\log(|G|/|G_{1234}|)}.$$

Note that Ingleton fails iff $\sigma < 0$, and a lower score means a larger violation. Essentially this definition forms a ray starting from the origin and passing through the point in $\mathbb{R}^{2^4-1}$ corresponding to an entropy

vector, then finds its intersection with the hyperplane $h_{1234} = 1$ and computes $-\Delta_h$ for that point to measure the Ingleton violation. The best Ingleton score in the family $PGL(2, q)$ is attained when $q = 13$, with $\sigma = -0.0270$. In [23] many violations obtained have lower Ingleton scores, hence are more effective than $PGL(2, q)$. In [30] a conjecture concerning the lowest Ingleton score attainable by an arbitrary entropy vector is proposed, but has been refuted recently by Matúš and Csirmaz [31].

A perhaps more geometrically meaningful scaling factor is the 2-norm of the entropy vector, as proposed in [32]:

Definition 4: For an entropy vector $h = (h_\alpha : \emptyset \neq \alpha \subseteq \{1, 2, 3, 4\})$, define the *Ingleton violation index* to be

$$\iota(h) = \frac{\Delta_h}{\|h\|_2} = \frac{\Delta_h}{\sqrt{h^T h}}.$$

Essentially this definition measures the "sine" of the angle between an entropy vector and the Ingleton hyperplane $\Delta_h = 0$. The Ingleton inequality fails iff $\iota > 0$, and a larger index means a larger violation. Note that two entropy vectors might have the same violation index but different Ingleton scores, and vice versa. The best Ingleton violation index in the family $PGL(2, q)$ is again attained when $q = 13$, with $\iota = 0.0082$, whereas for an arbitrary entropy vector the best ι found in literature is 0.0276 using quasi-uniform distributions [33].

Next we discuss two directions for generalizing the above Ingleton-violating family and finding new violations. On the one hand, $PGL(2, q)$ is the quotient group of $GL(2, q)$, so supposedly $GL(2, q)$ should have a richer choice of subgroups violating Ingleton inequality. This approach is explored in the next section. On the other hand, since the subgroups in the $PGL(2, q)$ family have simple but fundamental interpretations in terms of group actions, we can generalize them in this framework. In particular, we obtain two new families of violations in $PGL(n, q)$ for general n, and further generalize to an abstract construction using 2-transitive groups. Since this approach is more abstract and requires more background knowledge, we defer it to Section VII.

VI. Ingleton Violations in $GL(2, q)$

As $PGL(2, q)$ is the quotient group of $GL(2, q)$ modulo the subgroup V_q of scalar matrices, naturally one may ask if the general linear groups also violate Ingleton. In fact, the following lemma shows that there is at least one set of subgroups in $GL(2, q)$ that violates Ingleton for all finite field orders $q \geq 5$:

Lemma 2: If G is a finite group with a normal subgroup N such that $H \triangleq G/N$ has a set of Ingleton-violating subgroups, then the preimages of these subgroups under the natural homomorphism $g \mapsto gN$ are subgroups of G that also violate Ingleton.

Proof: Let $(H_i : 1 \leq i \leq 4)$ be a set of Ingleton-violating subgroups in H. Define G_i to be the preimage of H_i under the natural homomorphism, then G_i is a group containing N for each i. By the Lattice Isomorphism Theorem (see e.g. [26]), for any nonempty subset $\alpha \subseteq \{1, 2, 3, 4\}$, $G_\alpha/N = H_\alpha$, and so $|G_\alpha| = |H_\alpha| \cdot |N|$. Thus by checking the orders in (4), $(G_i : 1 \leq i \leq 4)$ also violate Ingleton. ■

Searching with GAP, we find $GL(2, 5)$ to be the smallest general linear group that violates Ingleton. Up to subscript symmetries and conjugations, it has 15 sets of Ingleton-violating subgroups. We would like to analyze their structures and generalize them for $q \geq 5$ if possible.

Throughout this section, we always assume q is a finite field order, and p is the characteristic of $\mathbb{F}_q$. We begin our analysis by identifying the preimages of the Ingleton-violating subgroups in the previous section under the natural homomorphism

$$\pi : GL(2, q) \to GL(2, q)/V_q = PGL(2, q),$$

according to Lemma 2. With no surprise, when $q = 5$ these correspond to one of the 15 violation instances in $GL(2, 5)$, and they take on nice matrix structures similar to the subgroups in Section V. Based on this set of subgroups we have 10 other instances, all of which are essentially its variants: each instance differs from the preimages at exactly one subgroup (either G_1 or G_2). These 11 violation instances can be easily extended to families of Ingleton-violating subgroups in $GL(2, q)$ for $q \geq 5$, sometimes with an extra condition. The remaining 4 instances cannot be derived directly from the preimages; however, they are interrelated and all their subgroups are equal or conjugate to some known subgroups from the previous instances. They also generalize to Ingleton-violating families in $GL(2, q)$ with some extra conditions.

Table II summarizes how the generalization of these instances depends on the values of p and q. We can see that when $p = 2$, these 15 instances collapse to only 6 dinstinct ones; also some instances need specific conditions on p and q to violate Ingelton.

In Table III, the orders of the subgroups for the cases we have explored in $PGL(2, q)$ and $GL(2, q)$ are listed. No. 0 denotes the instance in $PGL(2, q)$, and No. 1–15 denote the generalizations of the 15 violation instances in $GL(2, 5)$ to $GL(2, q)$. Since all instances have the subgroup order symmetries

$$|G_3| = |G_4|, \quad |G_{123}| = |G_{124}|, \quad |G_{13}| = |G_{14}|, \quad |G_{23}| = |G_{24}|,$$

only one of each pair of orders is listed. Note that when $p = 2$, there are only 6 such dinstinct generalizations, which are Instances 1, 2, 6, 7, 12 and 14. Thus for the order calculation of all other

TABLE II

(A) IDENTICAL INSTANCES WHEN $p = 2$ (B) CASES WHEN INGLETON IS NOT VIOLATED

Instance No.	Identical Instance(s)
1	5
2	3, 4
6	8, 10
7	9, 11
12	13
14	15

Instance No.	$p = 3$	$p \neq 2$, $\frac{q-1}{2}$ odd
8, 9		$\times$
12, 14	$\times$	
13, 15	$\times$	$\times$

TABLE III

ORDERS OF SUBGROUPS AND INTERSECTIONS

Ins. No.	$\lvert G_1\rvert$	$\lvert G_2\rvert$	$\lvert G_3\rvert$	$\lvert G_{34}\rvert$	$\lvert G_{123}\rvert$	$\lvert G_{12}\rvert$	$\lvert G_{13}\rvert$	$\lvert G_{23}\rvert$	$LHS - RHS$ in (4)
0	6	$q(q-1)$	$2(q-1)$	1	1	2	2	$q-1$	$2(q-1)(4-q)$
1	$6(q-1)$	$q(q-1)^2$	$2(q-1)^2$	$q-1$	$q-1$	$2(q-1)$	$2(q-1)$	$(q-1)^2$	$2(q-1)^6(4-q)$
2,4	6	$q(q-1)^2$	$2(q-1)^2$	$q-1$	1	2	2	$(q-1)^2$	$2(q-1)^3(4-q)$
3	12	$q(q-1)^2$	$2(q-1)^2$	$q-1$	2	4	4	$(q-1)^2$	$16(q-1)^3(4-q)$
5	$3(q-1)$	$q(q-1)^2$	$2(q-1)^2$	$q-1$	$\frac{q-1}{2}$	$q-1$	$q-1$	$(q-1)^2$	$\frac{1}{4}(q-1)^6(4-q)$
6–9	$6(q-1)$	$q(q-1)$	$2(q-1)^2$	$q-1$	1	2	$2(q-1)$	$q-1$	$2(q-1)^3(4-q)$
10,11	$6(q-1)$	$2q(q-1)$	$2(q-1)^2$	$q-1$	2	4	$2(q-1)$	$2(q-1)$	$16(q-1)^3(4-q)$
12–15	6	$q(q-1)$	$q(q-1)$	1	1	2	2	$q-1$	$2(q-1)(4-q)$
8',9'	$6(q-1)$	$q(q-1)$	$2(q-1)^2$	$q-1$	2	2	$2(q-1)$	$q-1$	$8(q-1)^3(2q+1)$
13',15'	6	$q(q-1)$	$q(q-1)$	2	1	1	1	$q-1$	$(q-1)(11q+1)$

instances in $GL(2,q)$ assume $p \neq 2$. Moreover, No. 8', 9', 13' and 15' correspond to Instances 8, 9, 13 and 15 when $p \neq 2$ but $\frac{q-1}{2}$ is odd, in which case Ingleton is satisfied. Finally, the order calculation for Instances 12–15 only works for $p \neq 3$. From Table III, we can calculate that all violation instances in the table have the same Ingleton ratio $r = 4(q-1)/(3q)$, which is the same as the family $PGL(2,q)$. But the scaling factors for both the Ingleton score and the violation index are no larger than $PGL(2,q)$ in these instances, so they are no more effective.

In the following, we present all of these extended violation families, with Section VI-A being the set of preimage subgroups, Sections VI-B and VI-C its 10 variants, and Section VI-D the remaining 4 instances. We continue to use the notations from Section V with t being a primitive element of $\mathbb{F}_q$, but

we redefine

$$N = \{\, A_\alpha |\, \alpha \in \mathbb{F}_q\} = \langle A_{\xi_1}, A_{\xi_2}, \ldots, A_{\xi_m}\rangle \cong \langle A_{\xi_1}\rangle \times \langle A_{\xi_2}\rangle \times \ldots \times \langle A_{\xi_m}\rangle \cong \mathbb{Z}_p^m.$$

In addition, we introduce the following matrices and subgroups in $GL(2,q)$ to facilitate our presentation:

$$B' = \begin{bmatrix} -1 & 0 \\ 0 & t \end{bmatrix}, \qquad P = \begin{bmatrix} t & 0 \\ 0 & 1 \end{bmatrix}, \qquad P' = \begin{bmatrix} t & 0 \\ 0 & -1 \end{bmatrix},$$

$$M = \langle C, B_1\rangle = \left\{ I, \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}, \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ -1 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & -1 \\ 0 & 1 \end{bmatrix} \right\},$$

$$K = \langle N, B\rangle = \left\{ \begin{bmatrix} 1 & 0 \\ \alpha & \beta \end{bmatrix} \middle| \begin{array}{l} \alpha \in \mathbb{F}_q, \\ \beta \in \mathbb{F}_q^\times \end{array} \right\}, \quad K' = \langle N, B'\rangle = \left\{ \begin{bmatrix} (-1)^k & 0 \\ \alpha & t^k \end{bmatrix} \middle| \begin{array}{l} \alpha \in \mathbb{F}_q, \\ k \in \mathcal{K}_q \end{array} \right\},$$

$$J = \langle N, P\rangle = \left\{ \begin{bmatrix} \beta & 0 \\ \alpha & 1 \end{bmatrix} \middle| \begin{array}{l} \alpha \in \mathbb{F}_q, \\ \beta \in \mathbb{F}_q^\times \end{array} \right\}, \quad J' = \langle N, P'\rangle = \left\{ \begin{bmatrix} t^k & 0 \\ \alpha & (-1)^k \end{bmatrix} \middle| \begin{array}{l} \alpha \in \mathbb{F}_q, \\ k \in \mathcal{K}_q \end{array} \right\}.$$

Note that when $p = 2$, we have $-1 = 1$, so $B' = B$, $P' = P$, and $K' = K$, $J' = J$. Also note that M and K precisely correspond to G_1 and G_2 in Section V, respectively. The group M is isomorphic to $D_6 \cong S_3$, while the other four groups are all semidirect products $\mathbb{Z}_p^m \rtimes \mathbb{Z}_{q-1}$, with $K \cong J$ and $K' \cong J'$. Moreover, K and J have generalized flower structures for all $q > 2$. However, if $p \neq 2$, K' and J' only have flower structures when $\frac{q-1}{2}$ is even, in which case they are also isomorphic to K. (See Section B-A in Appendices for proofs.) This turns out to be a necessary condition to violate Ingleton in all the instances where K' and J' are involved.

A. Instance 1: The Preimage Subgroups

To obtain the preimage H_0 of a subgroup $H \leq PGL(2,q)$ under π, we can generate H_0 in $GL(2,q)$ with the generators of H (without overlines) and tI, since $V_q = \langle tI\rangle \cong \mathbb{Z}_{q-1}$.

$G_1 = \langle tI, C, B_1\rangle = \langle V_q, M\rangle$. Since V_q is the center of $GL(2,q)$ and intersects M trivially, G_1 is a direct product: $G_1 = \{t^k X \,|\, X \in M, k \in \mathcal{K}_q\} \cong V_q \times M \cong \mathbb{Z}_{q-1} \times S_3$.

$G_2 = \langle tI, A_{\xi_1}, A_{\xi_2}, \ldots, A_{\xi_m}, B\rangle = \langle tI, N, B\rangle = \langle V_q, K\rangle$. G_2 is the subgroups of all lower triangular matrices in $GL(2,q)$, and as $V_q \bigcap K = 1$, we have $G_2 \cong V_q \times K \cong \mathbb{Z}_{q-1} \times (\mathbb{Z}_p^m \rtimes \mathbb{Z}_{q-1})$.

$G_3 = \langle tI, B_1C^2, B_3\rangle = \langle tI, CB_1, B_3\rangle = \langle CB_1, T\rangle$, where $T = \langle tI, B_3\rangle$. As $V_q \bigcap \langle B_3\rangle = 1$, we have $T = \{t^k B_3^m \,|\, k, m \in \mathcal{K}_q\} \cong V_q \times \langle B_3\rangle \cong \mathbb{Z}_{q-1} \times \mathbb{Z}_{q-1}$. It is easy to check that $(t^k B_3^m)^{CB_1} =$

TABLE IV

G_1 FOR INSTANCES 2–5

Ins. No.	2	3	4	5
G_1	$\langle C, B_1\rangle$	$\langle -C, B_1\rangle$	$\langle C, -B_1\rangle$	$\langle C, tB_1\rangle$

$t^{k+m}B_3^{-m} \in T$, so $G_3 = \langle CB_1\rangle \cdot T$ and $T \trianglelefteq G_3$. Furthermore, $|CB_1| = 2$ and $T \bigcap \langle CB_1\rangle = 1$, thus $G_3 \cong T \rtimes \langle CB_1\rangle \cong (\mathbb{Z}_{q-1} \times \mathbb{Z}_{q-1}) \rtimes \mathbb{Z}_2$ and

$$G_3 = \left\{ t^k \begin{bmatrix} 1 & 0 \\ t^m - 1 & t^m \end{bmatrix}, \quad t^{k+m} \begin{bmatrix} -1 & -1 \\ 1 - t^{-m} & 1 \end{bmatrix} \middle| \, k, m \in \mathcal{K}_q \right\}.$$

$G_4 = \langle tI, B_1C, B_4\rangle = \langle tI, B_1C, B\rangle = \langle B_1C, D\rangle$, where $D = \langle tI, B\rangle$. Since $V_q \bigcap \langle B\rangle = 1$, we have $D = \{t^k B^m \,|\, k, m \in \mathcal{K}_q\} = \{$all diagonal matrices in $GL(2, q)\} \cong V_q \times \langle B\rangle \cong \mathbb{Z}_{q-1} \times \mathbb{Z}_{q-1}$. Note that

$$\begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix}^{B_1C} = \begin{bmatrix} \beta & 0 \\ 0 & \alpha \end{bmatrix} \in D,$$

so $G_4 = \langle B_1C\rangle \cdot D$ and $D \trianglelefteq G_4$. Since $|B_1C| = 2$ and $D \bigcap \langle B_1C\rangle = 1$, $G_4 \cong D \rtimes \langle B_1C\rangle \cong (\mathbb{Z}_{q-1} \times \mathbb{Z}_{q-1}) \rtimes \mathbb{Z}_2$. Actually G_4 is the subgroups of all diagonal and anti-diagonal matrices in $GL(2, q)$:

$$G_4 = \left\{ \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix}, \begin{bmatrix} 0 & \beta \\ \alpha & 0 \end{bmatrix} \middle| \, \alpha, \beta \in \mathbb{F}_q^\times \right\}.$$

Calculating the intersections, we have $G_{12} = \langle tI, B_1\rangle \cong V_q \times \langle B_1\rangle$, $G_{13} = \langle tI, CB_1\rangle \cong V_q \times \langle CB_1\rangle$ and $G_{14} = \langle tI, B_1C\rangle \cong V_q \times \langle B_1C\rangle$, all of which are isomorphic to $\mathbb{Z}_{q-1} \times \mathbb{Z}_2$. Also, $G_{23} = T$, $G_{24} = D$ and $G_{34} = G_{123} = G_{124} = \langle tI\rangle = V_q$.

From the calculation in Table III, Ingleton is violated when $q \geq 5$.

B. Instances 2–5: Variants with Different G_1's

In all the instances in this section, only G_1 is different from Instance 1; it is now a *proper* subgroup of $\langle tI, C, B_1\rangle$ (see Table IV, where the generator-form for these groups is used to better demonstrate the subgroup relations). When $p \neq 2$, these instances are all distinct; however, when $p = 2$, clearly Instances 3 and 4 collapse to Instance 2, while Instance 5 becomes Instance 1. From Table III, we can see that they all violate Ingleton when $q \geq 5$.

1) Instance 2: $G_1 = M$.

$G_{12} = \langle B_1\rangle$, $G_{13} = \langle CB_1\rangle$ and $G_{14} = \langle B_1C\rangle$ are all isomorphic to $\mathbb{Z}_2$, and $G_{123} = G_{124} = 1$.

2) Instance 3: $G_1 = \langle -C, B_1 \rangle$.

We only consider the case $p \neq 2$, since otherwise this is the same as Instance 2. As $|C| = 3$, we have $(-C)^3 = -I$ and $(-C)^4 = C$. Thus $G_1 = \langle -I, C, B_1 \rangle = \langle -I, M \rangle \cong \langle -I \rangle \times M \cong \mathbb{Z}_2 \times S_3 \cong D_{12}$, since $\langle -I \rangle$ is a subgroup of V_q and intersects M trivially. So $G_1 = \{\pm X \,|\, X \in M\}$.

Now $G_{12} = \langle -I, B_1 \rangle \cong \langle -I \rangle \times \langle B_1 \rangle$, $G_{13} = \langle -I, CB_1 \rangle \cong \langle -I \rangle \times \langle CB_1 \rangle$ and $G_{14} = \langle -I, B_1C \rangle \cong \langle -I \rangle \times \langle B_1C \rangle$, all of which are isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$. Furthermore, $G_{123} = G_{124} = \langle -I \rangle \cong \mathbb{Z}_2$.

3) Instance 4: $G_1 = \langle C, -B_1 \rangle$.

Here we also need only consider the case $p \neq 2$. Observe that $|C| = 3$, $|{-}B_1| = 2$ and $(C \cdot (-B_1))^2 = (CB_1)^2 = I$. This gives us $G_1 = \{I, C, C^2, -B_1, -B_1C, -CB_1\}$, so $G_1 \cong D_6 \cong S_3$.

For the intersections, we have $G_{12} = \langle -B_1 \rangle$, $G_{13} = \langle -CB_1 \rangle$ and $G_{14} = \langle -B_1C \rangle$ all isomorphic to $\mathbb{Z}_2$, and $G_{123} = G_{124} = 1$.

4) Instance 5: $G_1 = \langle C, tB_1 \rangle$.

When $p = 2$, q is even. Since $|B_1| = 2$ and $|t| = q - 1$, we have $(tB_1)^q = tI$ and $(tB_1)^{q-1} = B_1$. Thus $G_1 = \langle tI, C, B_1 \rangle$ and this instance is the same as Instance 1.

Now assume $p \neq 2$. As q is odd, $|tB_1| = q - 1$. When k is even, $(tB_1)^k = t^k I$ and so $C^{(tB_1)^k} = C$. Otherwise $(tB_1)^k = t^k B_1$, then $C^{(tB_1)^k} = B_1 C B_1 = C^{-1}$ since $(CB_1)^2 = I$. So $G_1 = \langle tB_1 \rangle \cdot \langle C \rangle$ and $\langle C \rangle \trianglelefteq G_1$. Furthermore, $\langle tB_1 \rangle \bigcap \langle C \rangle = 1$ and $|C| = 3$, thus $G_1 \cong \langle C \rangle \rtimes \langle tB_1 \rangle \cong \mathbb{Z}_3 \rtimes \mathbb{Z}_{q-1}$ and $G_1 = \{t^k I,\, t^k C,\, t^k C^2 \;\big|\; k \text{ even},\, k \in \mathcal{K}_q\} \bigcup \{t^k B_1,\, t^k B_1 C,\, t^k CB_1 \;\big|\; k \text{ odd},\, k \in \mathcal{K}_q\}$.

The intersections are: $G_{12} = \langle tB_1 \rangle$, $G_{13} = \langle tCB_1 \rangle$ and $G_{14} = \langle tB_1C \rangle$ are all isomorphic to $\mathbb{Z}_{q-1}$, and $G_{123} = G_{124} = \langle t^2 I \rangle \cong \mathbb{Z}_{\frac{q-1}{2}}$.

C. Instances 6–11: Variants with Different G_2's

In all the instances in this section, only G_2 is different from Instance 1; it is now a *proper* subgroup of $\langle tI, N, B \rangle$ (see Table V). It is easy to see that these instances are distinct when $p \neq 2$; otherwise Instances 8 and 10 collapse to Instance 6, while Instances 9 and 11 become Instance 7. Thus in the analysis of Instances 8–11, we assume $p \neq 2$. From Table III, Instances 6, 7, 10, 11 violate Ingleton whenever $q \geq 5$; however, if $p \neq 2$, Instances 8 and 9 only violate Ingleton when in addition $\frac{q-1}{2}$ is even. Please refer to Section B-B in Appendices for the calculation of subgroup intersections in Instances 8 and 9.

1) Instance 6: $G_2 = K$.

In this case, $G_{12} = \langle B_1 \rangle \cong \mathbb{Z}_2$ and $G_{123} = G_{124} = 1$. Also $G_{23} = \langle B_3 \rangle$ and $G_{24} = \langle B \rangle$, both of which are isomorphic to $\mathbb{Z}_{q-1}$.

TABLE V

G_2 FOR INSTANCES 6–11

Ins. No.	6	7	8	9	10	11
G_2	$\langle N, B\rangle$	$\langle N, P\rangle$	$\langle N, B'\rangle$	$\langle N, P'\rangle$	$\langle -I, N, B\rangle$	$\langle -I, N, P\rangle$

2) Instance 7: $G_2 = J$.

Here $G_{12} = \langle -B_1\rangle \cong \mathbb{Z}_2$, $G_{123} = G_{124} = 1$. Also, $G_{23} = \langle t^{-1}B_3\rangle$ and $G_{24} = \langle P\rangle$, both isomorphic to $\mathbb{Z}_{q-1}$.

3) Instance 8: $G_2 = K'$.

$$G_{12} = \begin{cases} \langle B_1\rangle \cong \mathbb{Z}_2 & \text{if } \frac{q-1}{2} \text{ is even} \\ \langle -I\rangle \cong \mathbb{Z}_2 & \text{otherwise} \end{cases}, \quad G_{123} = G_{124} = \begin{cases} 1 & \text{if } \frac{q-1}{2} \text{ is even} \\ \langle -I\rangle \cong \mathbb{Z}_2 & \text{otherwise} \end{cases}.$$

In this case, $G_{23} = \langle -B_3^{\frac{q+1}{2}}\rangle$ and $G_{24} = \langle B'\rangle$ are both isomorphic to $\mathbb{Z}_{q-1}$.

4) Instance 9: $G_2 = J'$.

$$G_{12} = \begin{cases} \langle -B_1\rangle \cong \mathbb{Z}_2 & \text{if } \frac{q-1}{2} \text{ is even} \\ \langle -I\rangle \cong \mathbb{Z}_2 & \text{otherwise} \end{cases}, \quad G_{123} = G_{124} = \begin{cases} 1 & \text{if } \frac{q-1}{2} \text{ is even} \\ \langle -I\rangle \cong \mathbb{Z}_2 & \text{otherwise} \end{cases}.$$

Here $G_{23} = \langle tB_3^{\frac{q-3}{2}}\rangle$ and $G_{24} = \langle P'\rangle$ are isomorphic to $\mathbb{Z}_{q-1}$.

5) Instance 10: $G_2 = \langle -I, N, B\rangle$.

Now we have $G_2 = \langle -I, K\rangle \cong \langle -I\rangle \times K \cong \mathbb{Z}_2 \times (\mathbb{Z}_p^m \rtimes \mathbb{Z}_{q-1})$, since $\langle -I\rangle \bigcap K = 1$. Thus $G_2 = \{\pm X \mid X \in K\}$.

For the intersections, we have $G_{12} = \langle -I, B_1\rangle \cong \mathbb{Z}_2 \times \mathbb{Z}_2$ and $G_{123} = G_{124} = \langle -I\rangle \cong \mathbb{Z}_2$. Also, $G_{23} = \langle -I, B_3\rangle \cong \langle -I\rangle \times \langle B_3\rangle$ and $G_{24} = \langle -I, B\rangle \cong \langle -I\rangle \times \langle B\rangle$, both isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_{q-1}$.

6) Instance 11: $G_2 = \langle -I, N, P\rangle$.

Here $G_2 = \langle -I, J\rangle \cong \langle -I\rangle \times J \cong \mathbb{Z}_2 \times (\mathbb{Z}_p^m \rtimes \mathbb{Z}_{q-1})$, since $\langle -I\rangle \bigcap J = 1$. Thus $G_2 = \{\pm X \mid X \in J\}$.

Moreover, $G_{12} = \langle -I, -B_1\rangle = \langle -I, B_1\rangle \cong \mathbb{Z}_2 \times \mathbb{Z}_2$ and $G_{123} = G_{124} = \langle -I\rangle \cong \mathbb{Z}_2$. Also, $G_{23} = \langle -I, t^{-1}B_3\rangle \cong \langle -I\rangle \times \langle t^{-1}B_3\rangle$ and $G_{24} = \langle -I, P\rangle \cong \langle -I\rangle \times \langle P\rangle$ are both isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_{q-1}$.

D. Instances 12–15

For these last four instances, G_1 is always M, G_2–G_4 are equal or conjugate to one of K, K', J, J', as listed in Table VI. Thus G_2–G_4 are all semidirect products $\mathbb{Z}_p^m \rtimes \mathbb{Z}_{q-1}$ and the structures of G_3 and G_4

TABLE VI

SUBGROUPS FOR INSTANCES 12–15

Ins. No.	G_1	G_2	G_3	G_4
12	M	$\langle N, B\rangle$	$\langle N, P\rangle^E$	$\langle N, P\rangle^Q$
13	M	$\langle N, B'\rangle$	$\langle N, P'\rangle^E$	$\langle N, P'\rangle^Q$
14	M	$\langle N, P\rangle^E$	$\langle N, B\rangle$	$\langle N, B\rangle^W$
15	M	$\langle N, P'\rangle^E$	$\langle N, B'\rangle$	$\langle N, B'\rangle^W$

are different from all previous instances. The conjugators E, Q, W and the elements of new subgroups are listed as follows.

$$E = \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix}, \qquad Q = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}, \qquad W = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}.$$

$$J^E = \langle N, P\rangle^E = \left\{ \begin{bmatrix} 1-v & v \\ 1-u-v & u+v \end{bmatrix} \middle| \begin{array}{l} u \in \mathbb{F}_q^\times, \\ v \in \mathbb{F}_q \end{array} \right\},$$

$$(J')^E = \langle N, P'\rangle^E = \left\{ \begin{bmatrix} (-1)^j - \alpha & \alpha \\ (-1)^j - t^j - \alpha & t^j + \alpha \end{bmatrix} \middle| \begin{array}{l} \alpha \in \mathbb{F}_q, \\ j \in \mathcal{K}_q \end{array} \right\},$$

$$J^Q = \langle N, P\rangle^Q = \left\{ \begin{bmatrix} 1+2y & y \\ 2(x-2y-1) & x-2y \end{bmatrix} \middle| \begin{array}{l} x \in \mathbb{F}_q^\times, \\ y \in \mathbb{F}_q \end{array} \right\},$$

$$(J')^Q = \langle N, P'\rangle^Q = \left\{ \begin{bmatrix} (-1)^i + 2\beta & \beta \\ 2\left(t^i - 2\beta - (-1)^i\right) & t^i - 2\beta \end{bmatrix} \middle| \begin{array}{l} \beta \in \mathbb{F}_q, \\ i \in \mathcal{K}_q \end{array} \right\},$$

$$K^W = \langle N, B\rangle^W = \left\{ \begin{bmatrix} x & y \\ 0 & 1 \end{bmatrix} \middle| \begin{array}{l} x \in \mathbb{F}_q^\times, \\ y \in \mathbb{F}_q \end{array} \right\} = \left\{ X^T \mid X \in J \right\},$$

$$(K')^W = \langle N, B'\rangle^W = \left\{ \begin{bmatrix} t^i & \beta \\ 0 & (-1)^i \end{bmatrix} \middle| \begin{array}{l} \beta \in \mathbb{F}_q, \\ i \in \mathcal{K}_q \end{array} \right\} = \left\{ X^T \mid X \in J' \right\}.$$

As mentioned in Table II, Instances 12–15 do not violate Ingleton when $p = 3$. The reasons are as follows. If $p = 3$, then $2 = -1$, so $E = Q$ and $M \leq J^E$. Thus in Instance 12 we have $G_3 = G_4$ and $G_1 \leq G_3$, while in Instances 13 and 14 we have $G_3 = G_4$ and $G_1 \leq G_2$ respectively. So these three instances satisfy Conditions 5 and/or 7. Instance 15, however, satisfies Condition 3 in this case (see Section B-C in Appendices).

Besides, we also need $p \neq 2$ to make Instances 13 and 15 distinct: otherwise they collapse to Instances 12 and 14 respectively. Thus in the rest of this section, we always assume $p \neq 3$, while for Instances 13 and 15 we assume $p > 3$. From Table III, Instances 12 and 14 violate Ingleton when $q \geq 5$ (and of course, $p \neq 3$), while if $p \neq 2$, Instances 13 and 15 only violate Ingleton when in addition $\frac{q-1}{2}$ is even. Please refer to Section B-D in Appendices for the intersection calculations.

1) Instance 12: $G_2 = K, G_3 = J^E, G_4 = J^Q$.

We have $G_{12} = \langle B_1 \rangle$, $G_{13} = \langle B_1 C \rangle$ and $G_{14} = \langle C B_1 \rangle$ all isomorphic to $\mathbb{Z}_2$, and $G_{34} = G_{123} = G_{124} = 1$. Furthermore,

$$G_{23} = \left\{ \begin{bmatrix} 1 & 0 \\ 1-t^j & t^j \end{bmatrix} \middle| j \in \mathcal{K}_q \right\} = \langle P \rangle^E, \quad G_{24} = \left\{ \begin{bmatrix} 1 & 0 \\ 2(t^i-1) & t^i \end{bmatrix} \middle| i \in \mathcal{K}_q \right\} = \langle P \rangle^Q$$

both are isomorphic to $\mathbb{Z}_{q-1}$.

2) Instance 13: $G_2 = K', G_3 = (J')^E, G_4 = (J')^Q$.

When $\frac{q-1}{2}$ is even, G_{12}, G_{13}, G_{14} and G_{34} are the same as in Instance 12. Otherwise $G_{12} = G_{13} = G_{14} = 1$ and $G_{34} = \langle -I \rangle \cong \mathbb{Z}_2$. G_{123} and G_{124} are always trivial. Also,

$$G_{23} = \left\{ \begin{bmatrix} (-1)^j & 0 \\ (-1)^j - t^j & t^j \end{bmatrix} \middle| j \in \mathcal{K}_q \right\} = \langle P' \rangle^E, \quad G_{24} = \left\{ \begin{bmatrix} (-1)^i & 0 \\ 2\left(t^i - (-1)^i\right) & t^i \end{bmatrix} \middle| i \in \mathcal{K}_q \right\} = \langle P' \rangle^Q$$

are both isomorphic to $\mathbb{Z}_{q-1}$.

3) Instance 14: $G_2 = J^E, G_3 = K, G_4 = K^W$.

Observe that G_2 and G_3 are obtained from swapping the corresponding subgroups from Instance 12. Therefore G_{12} and G_{13} are also swapped while G_{23} remains the same. It turns out that G_{14}, G_{34}, G_{123} and G_{124} are also the same as in Instance 12. Furthermore,

$$G_{24} = \left\{ \begin{bmatrix} t^i & 1-t^i \\ 0 & 1 \end{bmatrix} \middle| i \in \mathcal{K}_q \right\} = \langle B \rangle^W \cong \mathbb{Z}_{q-1}.$$

4) Instance 15: $G_2 = (J')^E, G_3 = K', G_4 = (K')^W$.

In this case, G_2 and G_3 from Instance 13 are swapped to yield the corresponding subgroups here. So G_{12} and G_{13} are also swapped while G_{23} stays the same. Moreover, G_{14}, G_{34}, G_{123} and G_{124} are the same as in Instance 13, both when $\frac{q-1}{2}$ is even and otherwise. Finally,

$$G_{24} = \left\{ \begin{bmatrix} t^i & (-1)^i - t^i \\ 0 & (-1)^i \end{bmatrix} \middle| i \in \mathcal{K}_q \right\} = \langle B' \rangle^W \cong \mathbb{Z}_{q-1}.$$

VII. Interpretation and Generalizations of Violation in $PGL(2,q)$ using Theory of Group Actions

Instead of invertible matrices, we can also regard a general linear group as the group of all invertible linear transformations on a vector space. In this section, we take this point of view and consider the actions of linear groups on their corresponding projective geometries. Such actions induce a permutation representation for each general linear group on its projective geometry, and the projective linear groups are naturally defined in this framework. Using the theory of group actions, we show that the Ingleton violation in $PGL(2,q)$ from Section V has a nice interpretation: each subgroup is some sort of stabilizer for a set of points in the projective geometry. Furthermore, based on this understanding, we generalize the construction in $PGL(2,q)$ to two new families of Ingleton violations in $PGL(n,q)$ for a general n.[9] Finally, we provide an abstract construction in 2-transitive groups generalizing these ideas.

Throughout this section we assume basic knowledge in the theory of group actions, which can be found in standard group theory textbooks. In particular, we make extensive use of the orbit-stabilizer theorem, which says the order of the orbit of an element is equal to the index of it stabilizer (see e.g. [26, Sec. 4.1, Prop. 2]). Most notations are standard abstract algebra notations, see e.g. [26]; the rest are introduced when they first appear. Note that this section is more abstract than the others and assumes more background knowledge in abstract algebra.

This section is mostly based on Prof. M. Aschbacher's correspondences with us. We have furnished various details and explanations for clarity.

A. Preliminaries for Linear Groups

Let V be an n-dimensional vector space over a field F. Recall $GL(V)$ and $SL(V)$ are the general linear group and special linear group on V, respectively. They are examples of groups of Lie type, a notion which is not totally well defined.

Each group G of Lie type possesses a *building*, a simplicial complex on which G is represented as a group of automorphisms. A (abstract) *simplicial complex* consists of a set X of *vertices* together with a collection of nonempty subsets of X called *simplices*; the only axiom says that each nonempty subset of a simplex is a simplex.

Example 1: Let X be a partially ordered set. The *order complex* of X is the simplicial complex with vertex set X and with the simplices the nonempty chains in the poset.

[9]Note that with Lemma 2, the families in $PGL(n,q)$ can also be easily extended to families of violations in $GL(n,q)$.

Example 2: The *projective geometry* $PG(V)$ of V is the poset of nonzero proper subspaces of V, partially ordered by inclusion. The building of $GL(V)$ and $SL(V)$ is the order complex of this poset. Of course $GL(V)$ permutes the subspaces of V, supplying a representation of $GL(V)$ on $PG(V)$ whose kernel is the subgroup of scalar maps. The images of $GL(V)$ and $SL(V)$ in the automorphism group $Aut(PG(V))$ are the *projective general linear group* $PGL(V)$ and *projective special linear group* $PSL(V)$. Write $GL(n, F)$, $SL(n, F)$, $PGL(n, F)$, $PSL(n, F)$ for the corresponding group when $\dim(V) = n$ and the field is F.

Example 3: Specialize to the case $n = 2$. Then $PG(V)$ consists of the *points* of V; i.e. the 1-dimensional subspaces of V. This is the so-called *projective line*. Let $\mathcal{X} = \{x_1, x_2\}$ be a basis of V. We regard the projective line as $\Omega = F \cup \{\infty\}$, where ∞ denotes Fx_1 and for $e \in F$, e denotes $F(ex_1 + x_2)$. Then given an invertible matrix

$$M(a, b, c, d) = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

in $GL(V)$, one can check that, subject to the identification of $PG(V)$ with Ω, $M(a, b, c, d)$ acts on Ω via

$$M(a, b, c, d) : x \mapsto \frac{ax + b}{cx + d},$$

where arithmetic involving ∞ is suitably interpreted; e.g. $(a\infty + b)/(c\infty + d) = a/c$ if $c \neq 0$ and ∞ if $c = 0$. So we can regard $PGL(V) = PGL(2, F)$ as the group of these projective linear maps $M(a, b, c, d)$, $ad - bc \neq 0$ on the projective line Ω.

The following result is well known and easy to prove:

Lemma 3: $PGL(2, F)$ is *sharply 3-transitive* on the projective line $PG(V)$. That is, $PGL(V)$ is transitive on ordered 3-tuples of distinct points, and only the identity fixes three points.

Next we introduce several types of subgroups for these linear groups.

A *Borel subgroup* of a group G of Lie type is the stabilizer of a maximal simplex in its building.

Example 4: A maximal simplex in $PG(V)$ is a flag $\tau = (0 < V_1 < \cdots < V_{n-1} < V)$, where $\dim(V_k) = k$. If we pick a basis $\mathcal{X} = \{x_1, ..., x_n\}$ for V such that $V_k = \langle x_i : 1 \leq i \leq k \rangle$, then the Borel subgroup stabilizing τ is the subgroup whose matrices with respect to $\mathcal{X}$ are the upper triangular invertible matrices.

Let $G = PGL(2, F)$. By definition, the stabilizers $G_{Fx_1} = G_\infty$ and $G_{Fx_2} = G_0$ are both Borel subgroups of G. The matrices of these subgroups are upper triangular and lower triangular respectively. As G is transitive on Ω, for each of $u = \infty, 0$ we have the bijection $gG_u \mapsto g(u)$ of the coset space G/G_u with Ω (by orbit-stabilizer theorem).

Buildings have certain special subcomplexes called *apartments*. For a group G of Lie type, the pointwise stabilizer of an apartment is called a *Cartan subgroup* of G.

Example 5: In the projective geometry, the apartments are of the form $\Sigma(\mathcal{X})$ for $\mathcal{X} = \{x_1, \cdots, x_n\}$ a basis for V, where $\Sigma(\mathcal{X})$ consists of the subspaces spanned by nonempty proper subsets of $\mathcal{X}$. The matrices in the Cartan subgroup stabilizing $\Sigma(\mathcal{X})$ are the diagonal matrices.

Suppose $n = 2$. Then $\Sigma(\mathcal{X}) = \{Fx_1, Fx_2\} = \{\infty, 0\}$ is just a pair of points. The *global stabilizer* $G(u, v)$ of a pair of points is the subgroup of G permuting the 2-subset $\{u, v\}$. In $G = PGL(2, F)$ it is (usually) the normalizer of the Cartan subgroup and dihedral. Furthermore, $G_0 \cap G(0, \infty) = G_{0,\infty}$ is a Cartan subgroup isomorphic to the multiplicative group $F^\times$ of F.

Let G be $GL(V)$ or $PGL(V)$ in the rest of this section.

An element of $GL(V)$ is *unipotent* if all its eigenvalues are 1. A subgroup of $GL(V)$ is *unipotent* if all its elements are unipotent. The *unipotent radical* $Q(H)$ of a subgroup H of $GL(V)$ is the largest normal unipotent subgroup of H. For example if F is finite of characteristic p, then $Q(H)$ is the largest normal p-subgroup of H. Passing to images in $PGL(V)$, we have the corresponding notions in that group also.

A subgroup H of G is a *parabolic* if H is the stabilizer of a simplex in the projective geometry $PG(V)$. Thus for example Borel subgroups are parabolics, and indeed the parabolics are the overgroups of the Borel subgroups.

Example 6: Let $F = \mathbb{F}_q$, U an m-dimensional subspace of V with $0 < m < n$, $G = GL(V)$, and $H = N_G(U)$ the (global) stabilizer of U in G. As $\{U\}$ is a simplex in $PG(V)$, H is a parabolic. Pick a complement W to U in V, and let $\mathcal{X}_1$ and $\mathcal{X}_2$ be bases for U and W respectively. Then the matrices of H with respect to $\mathcal{X}_1 \cup \mathcal{X}_2$ have the form $\begin{bmatrix} K & L \\ 0 & R \end{bmatrix}$ with K and R invertible. Define

$$q_n = q^{n(n-1)/2}, \quad M_k = \prod_{i=1}^{k} (q^i - 1)$$

for $1 \leq k \leq n$, then

$$|GL(k, q)| = q_k M_k,$$

$$|H| = |GL(m, q)| \cdot |GL(n - m, q)| \cdot q^{m(n-m)} = q_n M_m M_{n-m}.$$

Furthermore, in $PGL(V)$ the image of H has order $q_n M_m M_{n-m}/(q - 1)$.

B. Interpretation of the Ingleton Violation in $PGL(2, q)$

Let $F = \mathbb{F}_q$ and $G = PGL(2, q) = PGL(2, \mathbb{F}_q)$. In the Ingleton violation construction in Section V we have a 4-tuple of subgroups $\rho = (G_i : 1 \leq i \leq 4)$ of G. The group $G_2 = G_{Fx_2} = G_0$ is a Borel

subgroup. The subgroups G_3 and G_4 are isomorphic to the dihedral group $D_{2(q-1)}$ of order $2(q-1)$, and their intersection G_{2i} with G_2 is cyclic of order $q-1$ and with G_{34} of order 1. This forces G_{2i}, $i = 3, 4$, to be distinct Cartan subgroups of G_2, and hence $G_i = G(0, e_i)$ for some $e_i \in F$. In fact from the forms of the matrices in G_3 and G_4 it is easy to check that $e_3 = -1$ and $e_4 = \infty$.

Finally $G_1 \cong S_3$ with G_{1i} being the three subgroups of G_1 of order 2 for $2 \leq i \leq 4$. For $2 \leq i \leq 4$ let $G_{1i} = \langle t_i \rangle$, and for $1 \leq j \leq 4$ let Δ_j be the orbit of G_j on Ω containing 0. Then $|\Delta_j| = |G_j : G_{2j}| = n_j$ where $n_3 = n_4 = 2$ and $n_1 = 3$. Indeed $\Delta_i = \{0, t_i(0)\}$ for $i = 3, 4$, with $\Delta_3 = \{0, -1\}$ and $\Delta_4 = \{0, \infty\}$. Then as $G_1 = \langle t_3, t_4 \rangle$ and $n_1 = 3$, $\Delta = \Delta_1 = \{0, -1, \infty\}$. But as G is sharply 3-transitive, the global stabilizer $G(\Delta)$ is isomorphic to S_3. Hence $G_1 = G(\Delta)$, and is determined by G_2, G_3 and G_4.

Hence the 4-tuple ρ is determined by the ordered triple $(0, -1, \infty)$ with the four subgroups being various (global) stabilizers on it. Furthermore, given an arbitrary ordered triple (α, β, γ) of distinct points in Ω, we can construct a 4-tuple ρ' in the same fashion, where $G_2 = G_\alpha$, $G_3 = G(\alpha, \beta)$, $G_4 = G(\alpha, \gamma)$, and $G_1 = G(\alpha, \beta, \gamma)$. Since G is 3-transitive on Ω, by the same element in G all four subgroups in ρ' are conjugate to their counterparts in ρ. In particular, the new tuple ρ' also violates Ingleton.

With respect to the "flower structure" of $G_2 = G_0$, this follows from the fact that G_0 is a Frobenius group on $\Omega' = \Omega - \{0\}$. That is, G_0 is a transitive permutation group on Ω' in which the maximum number of fixed points of a nonidentity element is 1. (This is guaranteed by the sharp 3-transitivity of G.) Then by a theorem of Frobenius, the identity 1 of G_0, together with the set of elements with no fixed points, forms a normal subgroup K called the *Frobenius kernel* of the Frobenius group. In our case, K is the subgroup N in Sections IV and V, which is the unipotent radical of the Borel subgroup G_0 and is isomorphic to the additive group of the field F. Also $G_0 - K$ is partitioned by the sets $G_{0,a} - \{1\}, a \in \Omega'$; these are the $|\Omega'| = q$ petals in the flower. The subgroups $G_{0,a}$ are the q Cartan subgroups contained in G_0, and each is isomorphic to $F^\times$.

C. Generalizations in $PGL(n, q)$

Let $\tau = (G_i : 1 \leq i \leq 4)$ be a family of subgroups of a finite group G. The Ingleton inequality (4) fails iff

$$|G_1 G_2| < \frac{|G_{13} G_{23}||G_{14} G_{24}|}{|G_{34}|}.$$

In all constructions we will consider in this section, $G_i = G_{1i}G_{2i}$ for $i = 3, 4$ and $|G_3| = |G_4|$. Also $|G_1G_2| = |G_1 : G_{12}||G_2|$. Hence in such constructions Ingleton is violated iff

$$|G_1 : G_{12}||G_2| < \frac{|G_3|^2}{|G_{34}|}, \tag{14}$$

and the Ingleton ratio (13) becomes

$$r(\tau) = \frac{|G_3|^2}{|G_1 : G_{12}||G_2||G_{34}|}.$$

Now we explore three different approaches trying to extend the $PGL(2, q)$ family of violations ρ to $PGL(n, q)$.

1) Generalization 1: Let $G = PGL(n, q)$ with $n \geq 3$. It is easy to see that G is doubly transitive on the points of $PG(V)$ and transitive on triples of independent points. Let P_i, $2 \leq i \leq 4$, be independent points in V, $\Delta_i = \{P_2, P_i\}$ for $i = 3, 4$, and $\Delta = \{P_2, P_3, P_4\}$. Set $G_2 = N_G(P_2)$, $G_i = N_G(\Delta_i)$, $i = 3, 4$, and $G_1 = N_G(\Delta)$. Let $\tau = (G_i : 1 \leq i \leq 4)$.

Now G_2 is a parabolic and by Example 6,

$$|G_2| = q_n M_{n-1}. \tag{15}$$

Next $D = P_2 + P_3 + P_4$ is a 3-dimensional subspace of V, so by Example 6 again, $|N_G(D)| = q_n M_3 M_{n-3}/(q-1)$. Further through calculation of the preimages in $GL(n, q)$ we have

$$|N_G(D) : G_1| = \frac{|GL(3, q)|}{6(q-1)^3} = \frac{q^3 M_3}{6(q-1)^3},$$

since G_1 acts as the symmetric group on Δ of order 3, and for each pair of points there are $q-1$ different choices of mappings. So

$$|G_1| = \frac{|N_G(D)| \cdot 6(q-1)^3}{q^3 M_3} = \frac{6q_n M_{n-3}(q-1)^2}{q^3}. \tag{16}$$

As G_1 is transitive on Δ of order 3, $|G_1 : G_{12}| = 3$. Therefore

$$|G_1 : G_{12}||G_2| = 3|G_2| = 3q_n M_{n-1}. \tag{17}$$

Also for $i = 3, 4$, G_i and G_{1i} are both transitive on Δ_i of order 2, so $|G_i : G_{2i}| = |G_{1i} : G_{12i}| = 2$. Thus $|G_{1i}G_{2i}| = |G_{1i} : G_{12i}||G_{2i}| = |G_i|$ and $G_i = G_{1i}G_{2i}$ for $i = 3, 4$. Since G is doubly transitive on the points, G_3 is conjugate to G_4 and so $|G_3| = |G_4|$. Further $U = P_2 + P_3$ is a 2-dimensional subspace of V, so by Example 6, $|N_G(U)| = q_n M_2 M_{n-2}/(q-1)$. Also by calculating the preimages $|N_G(U) : G_3| = |GL(2, q)|/(2(q-1)^2) = qM_2/(2(q-1)^2)$, so

$$|G_3| = \frac{|N_G(U)| \cdot 2(q-1)^2}{qM_2} = \frac{2q_n M_{n-2}(q-1)}{q}. \tag{18}$$

Finally $G_{34} = G_\Delta$ is the pointwise stabilizer of Δ. Since G_1 is 3-transitive on Δ, $|G_1 : G_{34}| = 3! = 6$. So by (16):

$$|G_{34}| = \frac{q_n M_{n-3}(q-1)^2}{q^3}. \tag{19}$$

It follows from (17), (18), and (19) that (14) is satisfied iff

$$3q_n M_{n-1} < \frac{4q_n^2 M_{n-2}^2 (q-1)^2 \cdot q^3}{q^2 \cdot q_n M_{n-3}(q-1)^2} = 4q_n q M_{n-2}(q^{n-2}-1)$$

which holds iff $3(q^{n-1}-1) < 4q(q^{n-2}-1)$ iff

$$q^{n-1} - 4q + 3 > 0. \tag{20}$$

This inequality holds when $n \geq 4$ or $n = 3$ and $q \geq 4$.

Since G is transitive on all triples of independent points, all 4-tuples in this generalization are conjugate to each other.

The Ingleton ratio is

$$r(\tau) = \frac{4q_n^2 M_{n-2}^2 (q-1)^2 \cdot q^3}{q^2 \cdot 3q_n M_{n-1} \cdot q_n M_{n-3}(q-1)^2} = \frac{4q(q^{n-2}-1)}{3(q^{n-1}-1)},$$

which approaches $4/3$ for large q or n. Whereas in the original instance ρ, $r(\rho) = 4(q-1)/(3q)$, which has the same asymptotics. But the scaling factors for both the Ingleton score and the violation index are usually larger than $PGL(2,q)$, so in general τ is less effective in violating Ingleton.

2) Generalization 2: As usual let $F = \mathbb{F}_q$ and $G = PGL(n,q)$, with $n \geq 2$. Let P_i, $2 \leq i \leq 4$, be distinct but dependent points in V. Thus $P_i = Fx_i$, $i = 2,3$, for two independent vectors $x_2, x_3 \in V$, and $P_4 = Fx_4$, where $x_4 = ex_2 + x_3$ for some $e \in F$. Let U, Δ, Δ_i, $i = 3,4$, and G_i, $1 \leq i \leq 4$, be defined the same as in Generalization 1. Note that when $n = 2$ this is our original construction ρ.

From Generalization 1, $|G_2| = q_n M_{n-1}$ and $|N_G(U)| = q_n M_2 M_{n-2}/(q-1)$. Since U is a 2-dimensional subspace of V, $PGL(U)$ is sharply 3-transitive on the points of U by Lemma 3. Now as Δ is a set of three distinct points in U, its global stabilizer in $PGL(U)$ is isomorphic to S_3. Thus G_1 is 3-transitive on Δ. Observe that each vector in $\{x_i : 2 \leq i \leq 4\}$ is a unique linear combination of the other two, with both coefficients nonzero. Then fixing a permutation of $\{P_i : 2 \leq i \leq 4\}$, there are only $q-1$ linear transformations in $GL(U)$ that respect this permutation. Hence $|N_G(U) : G_1| = |GL(2,q)|/(6(q-1)) = qM_2/(6(q-1))$, and

$$|G_1| = \frac{|N_G(U)| \cdot 6(q-1)}{qM_2} = \frac{6q_n M_{n-2}}{q}. \tag{21}$$

G_1 is transitive on Δ, while for $i = 3,4$, G_i and G_{1i} are both transitive on Δ_i. G is doubly transitive on the points of $PG(V)$. Thus from arguments in Generalization 1 we have $|G_1 : G_{12}||G_2| = 3q_n M_{n-1}$,

$G_i = G_{1i}G_{2i}$ for $i = 3, 4$, and $|G_3| = |G_4|$. Also $|G_3| = 2q_nM_{n-2}(q-1)/q$. Since $G_{34} = G_\Delta$ is of index 6 in G_1, by (21):

$$|G_{34}| = \frac{q_nM_{n-2}}{q}.$$

Thus (14) is satisfied iff

$$3q_nM_{n-1} < \frac{4q_n^2M_{n-2}^2(q-1)^2 \cdot q}{q^2 \cdot q_nM_{n-2}} = \frac{4q_nM_{n-2}(q-1)^2}{q}$$

which holds iff $3q(q^{n-1}-1) < 4(q-1)^2$ iff

$$3q\sum_{i=0}^{n-2} q^i - 4q + 4 < 0. \tag{22}$$

When $n = 2$, this inequality holds iff $q > 4$. When $n > 2$, however, it always fails because $3q^2 - q + 4 > 0$ for all q.

Therefore, the original instance ρ is the only successful case in this construction, with Ingleton ratio $r(\rho) = 4(q-1)/(3q)$.

3) Generalization 3: Again take $G = PGL(n, q)$ with $n \geq 3$. Let U_2 be a point of V, U_i, $i = 3, 4$, distinct 2-dimensional subspaces of V with $U_3 \cap U_4 = U_2$, and $U_1 = U_3 + U_4$ the 3-dimensional subspace of V generated by U_3 and U_4. Set $G_i = N_G(U_i)$ for $1 \leq i \leq 4$, and $\lambda = (G_i : 1 \leq i \leq 4)$. Then all the G_i are parabolics with $|G_2| = q_nM_{n-1}$ from (15), $|G_3| = |G_4| = q_nM_2M_{n-2}/(q-1)$, and $|G_1| = q_nM_3M_{n-3}/(q-1)$. As G_1 is transitive on the $(q^3-1)/(q-1) = q^2+q+1$ points in U_1, $|G_1 : G_{12}| = q^2+q+1$, so

$$|G_1 : G_{12}||G_2| = (q^2+q+1)q_nM_{n-1}.$$

For $i = 3, 4$, G_i and G_{1i} are both transitive on the $(q^2-1)/(q-1) = q+1$ points in U_i, so $G_i = G_{1i}G_{2i}$ for $i = 3, 4$. Also G_{34} is the subgroup of G fixing U_2 and the points U_3/U_2 and U_4/U_2 of the quotient space U_1/U_2; in particular it is a subgroup of G_1. If we pick a basis $\mathcal{X}_1 = \{x_3, x_2, x_4\}$ for U_1 such that $U_2 = \langle x_2 \rangle$ and $U_i = \langle x_2, x_i \rangle$ for $i = 3, 4$, then elements of G_{34} correspond to the linear transformations in $GL(U_1)$ whose matrices with respect to $\mathcal{X}_1$ take the form

$$\begin{bmatrix} a & 0 & 0 \\ x & b & y \\ 0 & 0 & c \end{bmatrix},$$

where a, b and c are nonzero. So $|G_1 : G_{34}| = |GL(3, q)|/(q^2(q-1)^3) = qM_3/(q-1)^3$, and

$$|G_{34}| = \frac{|G_1|}{qM_3/(q-1)^3} = \frac{q_nM_3M_{n-3} \cdot (q-1)^3}{(q-1) \cdot qM_3} = \frac{q_nM_{n-3}(q-1)^2}{q}.$$

It follows that (14) is satisfied iff

$$(q^2+q+1)q_nM_{n-1} < \frac{q_n^2M_2^2M_{n-2}^2\cdot q}{(q-1)^2\cdot q_nM_{n-3}(q-1)^2} = q_nq(q+1)^2(q^{n-2}-1)M_{n-2},$$

which holds iff $(q^2+q+1)(q^{n-1}-1) < q(q+1)^2(q^{n-2}-1)$ iff

$$q^n - q^3 - q^2 + 1 > 0,$$

which holds iff $n \geq 4$.

The Ingleton ratio is

$$r(\lambda) = \frac{q_n^2M_2^2M_{n-2}^2\cdot q}{(q-1)^2\cdot(q^2+q+1)q_nM_{n-1}\cdot q_nM_{n-3}(q-1)^2} = \frac{q(q+1)^2(q^{n-2}-1)}{(q^2+q+1)(q^{n-1}-1)},$$

which approaches 1 for large q and $(q+1)^2/(q^2+q+1)$ (which is smaller than $4/3$) for large n. So this generalization seems less effective than the other two.

D. Generalizations in General 2-transitive Groups

In the following we generalize the Ingleton violation ρ in $PGL(2,q)$ to a more abstract construction, which includes Generalizations 1 and 2 as special cases.

Let G be a doubly transitive group on a set Ω of order $l \geq 3$, let α and β be distinct points in Ω, and assume $\gamma \in \Omega - \{\alpha,\beta\}$ such that the global stabilizer $G(\Delta)$ of $\Delta = \{\alpha,\beta,\gamma\}$ acts as the symmetric group on Δ (which is clearly the case when G is 3-transitive). Let $G_2 = G_\alpha$, $G_3 = G(\alpha,\beta)$, $G_4 = G(\alpha,\gamma)$, and $G_1 = G(\Delta)$. Set $\mu = (G_i : 1 \leq i \leq 4)$.

Let $k = \left|G_{\alpha,\beta}\right|$, $d = |G_\Delta|$, Γ the orbit of γ under the action of $G_{\alpha,\beta}$, and $c = |\Gamma|$. Observe that $c = |G_{\alpha,\beta} : G_\Delta| = k/d$ and $c \leq l-2$ as $\Gamma \subseteq \Omega - \{\alpha,\beta\}$. Further $c = l-2$ iff G is 3-transitive.

Since G is 2-transitive on Ω, G_2 is transitive on $\Omega - \{\alpha\}$ and so $|G_2 : G_{\alpha,\beta}| = l-1$. Also $|G_1 : G_{12}| = 3$ as G_1 is transitive on Δ, thus

$$|G_1 : G_{12}||G_2| = 3|G_2| = 3(l-1)k.$$

Next G_3 is conjugate to G_4 by 2-transitivity of G and for $i = 3,4$, G_i and G_{1i} are both transitive on Δ_i of order 2, so $G_{1i}G_{2i} = G_i$ and $|G_i| = 2k$ for $i = 3,4$. Finally $G_{34} = G_\Delta$ is of order d. Thus

$$|G_3|^2/|G_{34}| = 4k^2/d = 4kc,$$

so condition (14) is satisfied iff $3(l-1)k < 4kc$ iff

$$3(l-1) < 4c. \tag{23}$$

Further the Ingleton ratio $r(\mu) = 4c/(3(l-1))$.

If G is 3-transitive then $c = l-2$, so $3(l-1) < 4c = 4(l-2)$ iff $l > 5$. Further $r(\mu) = 4(l-2)/(3(l-1))$.

Both Generalization 1 and 2 fit in this construction, with ρ being the only 3-transitive case. In Generalization 1, $l = (q^n - 1)/(q - 1)$ and by independence of points in Δ,

$$c = \frac{(q^n - 1) - (q^2 - 1)}{q - 1} = \frac{q^2(q^{n-2} - 1)}{q - 1},$$

so by (23), (14) is satisfied iff

$$3(\frac{q^n - 1}{q - 1} - 1) < \frac{4q^2(q^{n-2} - 1)}{q - 1},$$

which gives (20). In Generalization 2, l has the same value, but since $GL(U)$ is 3-transitive on the $(q^2 - 1)/(q - 1) = q + 1$ points of U, $c = q + 1 - 2 = q - 1$. Then by (23), (14) is satisfied iff

$$3(\frac{q^n - 1}{q - 1} - 1) < 4(q - 1),$$

which gives (22).

We see that the 3-transitive groups give rise to simple and effective Ingleton violation constructions. This category of groups include the alternating and symmetric groups, the groups $PGL(2, q)$ with $l = q + 1$, the Mathieu groups, the affine groups of degree 2^e (which are the semidirect product of an e-dimensional vector space E over $\mathbb{F}_2$ by $GL(E)$), and the subgroup of the affine group for $e = 4$ where the complement is A_7 rather than $GL(4, 2) \cong A_8$.

VIII. Considerations for Constructing Group Network Codes

We can use our Ingleton-violating groups to build group network codes. From Appendix A, the resulting entropy vectors are characterizable by the subgroups used, thus they are capable of violating the Ingleton inequality. In contrast, the entropy vectors of linear network codes always respect Ingleton. Furthermore, let G be any of $PGL(n, p)$, $PGL(n, q)$, $GL(n, p)$ or $GL(n, q)$. We will show in the following that linear network codes can be embedded in the group network codes constructed with direct products of copies of G. Apparently a direct product of any copies of an Ingleton-violating group still violates Ingleton, thus such classes of group network codes are strictly more powerful than linear network codes.

To construct a group network code, the choices of subgroups are not arbitrary: they should meet requirements (R1)–(R3). In particular, (R1) limits what subgroups can be associated with the sources: they need to satisfy

$$\prod_{s \in \mathcal{S}} |G_s| = |G|^{|\mathcal{S}|-1} |G_{\mathcal{S}}|. \tag{24}$$

When this is the case, we simply say the subgroups $\{G_s : s \in \mathcal{S}\}$ are independent in G. We will study the constructions of independent source subgroups in the context of $PGL(2, q)$ and $GL(2, q)$ (since they

have simpler structures than the other higher-degree linear groups), and also provide a universal source subgroup construction for direct products of groups.

A. Embeddings of Linear Network Codes

As discussed in Appendix A-C, linear network codes are a special type of group network codes. In particular, they are determined by the underlying additive group structure. The direct sum V of source vector spaces can be called the *ambient vector space* of a linear network code. Let $(V,+)$ denote the additive group of V. If we can find a finite group G such that $(V,+) \leq G$, then the linear network code is said to be *embedded* in the group network codes using G, since we can use subgroups of G to construct an equivalent group network code.

Consider a linear network code with ambient vector space $V = \mathbb{F}_q^d$ for some d and q, where $q = p^m$ for some prime p and some integer m. Observing that $\mathbb{F}_q$ is an m-dimensional vector space over $\mathbb{F}_p$, we can establish the following facts:

i) $(\mathbb{F}_p,+) \cong \mathbb{Z}_p$,

ii) $(\mathbb{F}_q,+) \cong (\mathbb{F}_p,+)^m \cong \mathbb{Z}_p^m$,

iii) $(V,+) \cong (\mathbb{F}_q,+)^d \cong \mathbb{Z}_p^{md}$.

Thus $(V,+)$ is embedded in the direct product of $m \cdot d$ copies of a group G, provided that G contains an element of order p—by Cauchy's theorem, this condition is equivalent to p divides $|G|$. It then follows that linear network codes over $\mathbb{F}_q$ are embedded in the group network codes using direct products of copies of G^m. In particular, let G be any of the linear groups $PGL(2,p)$, $PGL(2,q)$, $GL(2,p)$ or $GL(2,q)$. We have the following embeddings in these groups, using properties of the matrix A and the subgroup N:

1) In $PGL(2,p)$, $\left|\overline{A}\right| = p$. So $(V,+) \cong \left\langle \overline{A} \right\rangle^{md} \leq PGL(2,p)^{md}$.

2) In $GL(2,p)$, $|A| = p$. So $(V,+) \cong \langle A \rangle^{md} \leq GL(2,p)^{md}$.

3) In $PGL(2,q)$, $N = \left\{ \overline{A_\alpha} \,\middle|\, \alpha \in \mathbb{F}_q \right\} \cong \mathbb{Z}_p^m$. So $(V,+) \cong N^d \leq PGL(2,q)^d$.

4) In $GL(2,q)$, $N = \{ A_\alpha | \, \alpha \in \mathbb{F}_q \} \cong \mathbb{Z}_p^m$. So $(V,+) \cong N^d \leq GL(2,q)^d$.

Therefore, we also have the corresponding network code embeddings. Furthermore, these results for the degree-2 linear groups are easily extended to degree n, since the former are subgroups of the latter.

B. Sources Independence Requirement Considerations

If we want to utilize the Ingleton-violating groups $PGL(2,q)$ and $GL(2,q)$ to construct network codes, we need to find their independent subgroups. GAP searching shows that up to conjugation, $PGL(2,5)$

has 16 independent pairs of subgroups, 1 triple and no quadruple. For $GL(2,5)$, the numbers are 86, 14 and 0, respectively. It might be desirable to use some of the Ingleton-violating subgroups as sources, but we find no independent pairs in any violation instance in either $PGL(2,5)$ or $GL(2,5)$. Furthermore, we can prove the following negative results:

Lemma 4: Let $i,j \in \{1,2,3,4\}$ and $(i,j) \neq (3,4)$. For four random variables X_1, X_2, X_3 and X_4, if X_i and X_j are independent, then the Ingelton inequality (3) is satisfied.

Proof: By symmetry of (3), we only need to prove the result for when $(i,j) = (1,2)$ or $(1,3)$. In the first case, $h_{12} = h_1 + h_2$, so

$$\begin{aligned} h_{12} + h_{13} + h_{14} + h_{23} + h_{24} &\geq h_1 + h_2 + h_3 + h_{123} + h_4 + h_{124} \\ &\geq h_1 + h_2 + h_{34} + h_{123} + h_{124}, \end{aligned}$$

where we used $h_{13} + h_{23} \geq h_3 + h_{123}$ and $h_{14} + h_{24} \geq h_4 + h_{124}$ by submodularity of entropy. The second case is similar. ■

Corollary 1: There is no independent triple or quadruple in a set of four subgroups that violates (4).

On another note, if we want to use the Ingleton-violating subgroups in the network, Proposition 6 in Appendix A tells us that their intersection should contain the intersection of all the source subgroups. Since in $PGL(2,q)$ the intersection of the Ingleton-violating subgroups is trivial, we need to find trivially intersecting independent subgroups to serve as sources. In $PGL(2,5)$, there are 4 such pairs and no such triples. At least one of these pairs also extends to a general family:

Proposition 1: Let $U = \begin{bmatrix} 0 & -1 \\ t & 0 \end{bmatrix} \in GL(2,q)$, where t is a primitive element in $\mathbb{F}_q$. Let H be the image of $SL(2,q)$ in $PGL(2,q)$ under the natural homomorphism, which is isomorphic to $PSL(2,q)$. When $p \neq 2$, H and $\langle \overline{U} \rangle$ are independent in $PGL(2,q)$ with trivial intersection.

Proof: It is easy to see $|\overline{U}| = 2$, $\det U = t$. The determinant of any matrix representing an element in H takes the form $t^{2k} \in \langle t^2 \rangle$, for some k. But $t \notin \langle t^2 \rangle$ as $q-1$ is even, so $H \bigcap \langle \overline{U} \rangle = 1$. Also $|\langle \overline{U} \rangle| \cdot |H| = 2 \cdot |SL(2,q)|/2 = |SL(2,q)| = |PGL(2,q)|$, thus (24) holds. ■

In $GL(2,q)$ there are more Ingleton-violating instances, which have various intersections. So the requirement on the sources is not so strict and we have a richer class of subgroups to work with. As in $PGL(2,q)$, there exist trivially intersecting independent pairs, for example:

Proposition 2: In $GL(2,q)$, $SL(2,q)$ and $\langle B \rangle$ (or $\langle P \rangle$) are independent with trivial intersection.

Proof: Obviously $\det B^k = 1$ iff $B^k = I$, so $SL(2,q)$ and $\langle B \rangle$ have trivial intersection. Also $|B| \cdot |SL(2,q)| = (q-1) \cdot |GL(2,q)|/(q-1) = |GL(2,q)|$, thus (24) is satisfied. The proof for $\langle P \rangle$ is similar. ■

In general it is not easy to find many independent subgroups in a group. If the group is a direct product of n of its subgroups, however, it admits a natural construction of n independent subgroups:

Proposition 3: If $G = G_1 \times G_2 \times \cdots \times G_n$, then $1 \times G_2 \times \cdots \times G_n$, $G_1 \times 1 \times \cdots \times G_n$, ..., and $G_1 \times G_2 \times \cdots \times 1$ are n trivially intersecting independent subgroups in G.

Proof: Trivial intersection is obvious, and it is easy to check that both sides of (24) are equal to $\prod_{i=1}^{n} |G_i|^{n-1}$. ■

This construction is the generalization of the source construction for linear network codes, in which case the subgroup at source s is the W_s defined in Appendix A-C. Also we see that using direct products we can obtain independent subgroups for an arbitrary number of sources, but the group order also grows.

If we further require the sources to be of the same alphabet size, then the independent subgroups must have the same order. In the above proposition, this can be simply achieved by choosing G_i to be the same subgroup for each i. Additionally, for an arbitrary pair of independent subgroups, we have the following proposition.

Proposition 4: If G_s and G_r are independent in G, then $G_s \times G_r$ and $G_r \times G_s$ are independent in G^2 with the same order.

Proof: G_s and G_r satisfy $|G_s||G_r| = |G||G_s \bigcap G_r|$. Thus for the direct product construction, the LHS and RHS of (24) are $|G_s|^2|G_r|^2$ and $|G|^2|G_s \bigcap G_r|^2$ respectively, which are equal. ■

IX. Conclusion

Using a refined search we find the smallest group to violate the Ingleton inequality to be the 120 element group S_5. Investigating the detailed structure of the subgroups allowed us to determine that this is an instance of the Ingleton-violating family of groups $PGL(2, q)$ for prime powers $q \geq 5$. As this family has a nice interpretation in the theory of group actions, we generalize the idea to obtain more Ingleton violations in $PGL(n, q)$ and $GL(n, q)$. We also examine the preimage group $GL(2, q)$ of $PGL(2, q)$ and discover more families of violating subgroups. Nevertheless, even in $PGL(2, q)$ and $GL(2, q)$ for $q > 5$, there might still exist more violation instances that we have not explored, let alone other interesting groups. For example, subsequent to our work Boston and Nan [23] find many new violations in the class of permutation groups. Presumably there are infinite families of Ingleton violating groups, so the list of such families to date is by no means comprehensive and is far from complete.

The PGL and GL groups violate the Ingleton inequality and, since they contain linear network codes inside them, can provide network codes more powerful than linear ones. Developing group network codes requires designing the source subgroups that satisfy independence (R1) and the edge subgroups that satisfy

(R2) and (R3). The coding process requires two fundamental operations: (i) determining the intersection of all cosets from each incoming edge, and (ii) finding the appropriate coset for the outgoing overgroup of the intersected subgroups. Therefore constructing network codes from $PGL(n,q)$ and $GL(n,q)$ will require a thorough understanding of the structure of their subgroups and the corresponding coset operation. Investigating this issue may be a fruitful direction for future work.

APPENDIX A
GROUP NETWORK CODES: DETAILS

A. Code Construction

To establish the encoding and decoding process, we need an auxiliary lemma.

Lemma 5: Let K_1, K_2 be two subgroups of G with $K_1 \leq K_2$. Then the coset mapping

$$\begin{aligned} \pi : G/K_1 &\to G/K_2 \\ xK_1 &\mapsto xK_2 \end{aligned} \tag{25}$$

is a well defined onto function, where xK_1 is mapped to the unique coset in G/K_2 that contains it. Furthermore, if Λ_1 is a uniform random variable on G/K_1, then $\pi(\Lambda_1)$ is uniform on G/K_2.

Proof: π is well defined since $xK_2 = x'K_2$ whenever $xK_1 = x'K_1$. Note that K_2 is partitioned by the m distinct cosets $\{y_iK_1 : 1 \leq i \leq m\}$, where $m = |K_2/K_1|$ and $y_i \in K_2$ for $i = 1, 2, \ldots, m$. Therefore, each $xK_2 \in G/K_2$ is also partitioned by the m cosets $\{(xy_i)K_1 : 1 \leq i \leq m\}$, which are precisely the m preimages of xK_2 under π. Thus $\pi(\Lambda_1)$ is uniform on G/K_2. ■

For any collection α of subgroups of G, the intersection mapping (1) is a bijection. Consider the collection of all source subgroups. Let $\mathcal{X}_\mathcal{S} = \{(xG_s : s \in \mathcal{S}) \mid x \in G\} \subseteq \prod_{s\in\mathcal{S}} \mathcal{Y}_s$, then we have the bijective intersection mapping $\Theta_\mathcal{S} : \mathcal{X}_\mathcal{S} \to G/G_\mathcal{S}$. But with (R1), $\left|\prod_{s\in\mathcal{S}} \mathcal{Y}_s\right| = |G/G_\mathcal{S}| = |\mathcal{X}_\mathcal{S}|$ and so

$$\mathcal{X}_\mathcal{S} = \prod_{s\in\mathcal{S}} \mathcal{Y}_s.$$

This means that any coset tuple $(x_sG_s : s \in \mathcal{S})$ in $\prod_{s\in\mathcal{S}} \mathcal{Y}_s$ can be represented in the form $(xG_s : s \in \mathcal{S})$ for a common $x \in G$, and the intersection of $\{x_sG_s : s \in \mathcal{S}\}$ is equal to $xG_\mathcal{S}$. Therefore, we can rewrite the bijection $\Theta_\mathcal{S}$ as

$$\Theta_\mathcal{S} : \prod_{s\in\mathcal{S}} \mathcal{Y}_s \to G/G_\mathcal{S},$$

which maps a tuple to the intersection of all its cosets.

Moreover, let t be an edge or a sink node, define $\mathcal{X}_{\mathcal{I}(t)} = \{(xG_f : f \in \mathcal{I}(t)) \mid x \in G\}$ and $G_{\mathcal{I}(t)} = \bigcap_{f\in\mathcal{I}(t)} G_f$. Then the intersection mapping

$$\Theta_{\mathcal{I}(t)} : \mathcal{X}_{\mathcal{I}(t)} \to G/G_{\mathcal{I}(t)}$$

is a bijection. With (R2) and (R3), we can also define coset mappings for edges and source/sink pairs as follows. For each edge e, since $G_{\mathcal{I}(e)} \leq G_e$ by (R2), define the coset mapping π_e as (25) with $K_1 = G_{\mathcal{I}(e)}$ and $K_2 = G_e$. Similarly for each source s with $u \in \mathcal{D}(s)$, since $G_{\mathcal{I}(u)} \leq G_s$ by (R3), define $\pi_{u,s}$ with $K_1 = G_{\mathcal{I}(u)}$ and $K_2 = G_s$.

Now we can define the encoding and decoding functions. At each edge e, let the encoding function be $\phi_e = \pi_e \circ \Theta_{\mathcal{I}(e)}$. For each source s with $u \in \mathcal{D}(s)$, let the decoding function be $\phi_{u,s} = \pi_{u,s} \circ \Theta_{\mathcal{I}(u)}$. In other words, at an edge or a sink node t, the encoding/decoding function takes an input coset tuple $(Y_f : f \in \mathcal{I}(t))$ and first forms the intersection of them, which is a coset of $G_{\mathcal{I}(t)}$, then maps this coset to the unique coset of G_e (or G_s, whichever is appropriate) that contains it. Such network operations define a proper network code, since by the proposition below the decoding functions always yield correct source symbols at each sink node.

Proposition 5: Assume (R1) holds, and let the encoding and decoding functions be defined as above. Then for some common $x \in G$, $\forall s \in \mathcal{S}$, $Y_s = xG_s$ and $\forall e \in \mathcal{E}$, $Y_e = xG_e$. Also for each source s with $u \in \mathcal{D}(s)$, Y_s is recovered by the decoding function $\phi_{u,s}$.

Proof: Let the source symbols $(Y_s : s \in \mathcal{S})$ be an arbitrary tuple from $\prod_{s\in\mathcal{S}} \mathcal{Y}_s$. Since (R1) is true, as discussed above, for all $s \in \mathcal{S}$, $Y_s = xG_s$ with a common $x \in G$. As $\mathcal{G}$ is directed and acyclic, we can define the "depth" of each node v as the length of the longest path from a source node to v, and define the depth of an edge to be the depth of its tail node. Note that e is always "deeper" than f if $f \in \mathcal{I}(e)$. Also if $Y_f = xG_f$ for all $f \in \mathcal{I}(e)$, then $Y_e = \phi_e(Y_f : f \in \mathcal{I}(e)) = xG_e$. So by induction on the depths of the edges, $Y_e = xG_e$ for all $e \in \mathcal{E}$.

Furthermore, for each $s \in \mathcal{S}$ with $u \in \mathcal{D}(s)$, since $Y_f = xG_f$ for all $f \in \mathcal{I}(u)$, $\phi_{u,s}(Y_f : f \in \mathcal{I}(u)) = xG_s = Y_s$. Thus the source symbol Y_s is successfully recovered at u. ■

Remark 4: Note that the encoding/decoding function for an edge or a sink node t is only defined on $\mathcal{X}_{\mathcal{I}(t)}$, but not on the entire Cartesian product $\prod_{f\in\mathcal{I}(t)} \mathcal{Y}_f$. This is because for an arbitrary tuple in $\prod_{f\in\mathcal{I}(t)} \mathcal{Y}_f$, it is possible that the intersection of all cosets is the empty set, which is not a coset of $G_{\mathcal{I}(t)}$. However, with (R1) this is not a problem, as Proposition 5 guarantees that $(Y_f : f \in \mathcal{I}(t))$ is always a tuple in $\mathcal{X}_{\mathcal{I}(t)}$.

Remark 5: From the proof above, even without (R1) these encoding and decoding functions still constitute a valid network code, if the sources cooperate in such a way that the transmit tuples are always from $\mathcal{X}_{\mathcal{S}}$. But in this case the source random variables are dependent.

B. The Entropy Vector

Here we analyze the global mappings of this group network code, and show that the entropy vector is characterizable by the group G and its subgroups $\{G_t : t \in \mathcal{S} \cup \mathcal{E}\}$ when the sources are independent and uniform. First we give another auxiliary lemma.

Lemma 6: Let $K \leq G$ and let G_i, $i = 1, \ldots, n$, be subgroups of G containing K. For each i let π_i be the coset mapping defined as (25) with $K_1 = K$ and $K_2 = G_i$. Let Λ_K be a uniform random variable on G/K, and define $X_i = \pi_i(\Lambda_K)$ for each i. Then the entropy vector of $\{X_1, X_2, \ldots, X_n\}$ is exactly the group characterizable vector induced by G and $\{G_1, G_2, \ldots, G_n\}$.

Proof: For each nonempty subset $\alpha \subseteq \mathcal{N}$, since $K \leq G_\alpha$, we can define the coset mapping π_α with K and G_α. As in Section I-A, the alphabet of X_α is still $\mathcal{X}_\alpha = \{(xG_i : i \in \alpha) \mid x \in G\}$, and the intersection mapping Θ_α is a bijection. Also $\Theta_\alpha(X_\alpha) = \pi_\alpha(\Lambda_K)$, which is uniform on G/G_α by Lemma 5. So the joint entropy $H(X_\alpha) = H(\Theta_\alpha(X_\alpha)) = \log \frac{|G|}{|G_\alpha|}$ and the lemma follows. ■

For each $s \in \mathcal{S}$ define the coset mapping π'_s as (25) with $K_1 = G_\mathcal{S}$ and $K_2 = G_s$. For every edge e we can similarly define a new coset mapping π'_e with $K_1 = G_\mathcal{S}$ and $K_2 = G_e$, since according to the following proposition, $G_\mathcal{S} \leq G_e$.

Proposition 6: If (R2) is satisfied, then $\forall e \in \mathcal{E}$, $G_\mathcal{S} \leq G_e$.

Proof: The proposition is trivially true if e is emitted from a source node. Also if $G_\mathcal{S} \leq G_f$ for all $f \in \mathcal{I}(e)$, then by (R2) we have $G_\mathcal{S} \leq G_e$. Similar to Proposition 5, by induction on the depths of the edges the proof follows. ■

Proposition 7: $\forall e \in \mathcal{E}$, the global mapping at e for the above group network code is $\varphi_e = \pi'_e \circ \Theta_\mathcal{S}$. In other words, φ_e first forms the intersection of all the source cosets to obtain a coset of $G_\mathcal{S}$, and then maps this coset to the unique coset of G_e containing it.

Proof: Assume the source symbols $(Y_s : s \in \mathcal{S})$ are transmitted and let $\Lambda_\mathcal{S} = \Theta_\mathcal{S}(Y_s : s \in \mathcal{S})$. Then $\Lambda_\mathcal{S} = xG_\mathcal{S}$ for some $x \in G$, and $Y_s = xG_s = \pi'_s(\Lambda_\mathcal{S})$ for all $s \in \mathcal{S}$. By Proposition 5, $Y_e = xG_e = \pi'_e(\Lambda_\mathcal{S})$, so $\varphi_e = \pi'_e \circ \Theta_\mathcal{S}$. ■

Let the source random variables $\{Y_s : s \in \mathcal{S}\}$ be independent and uniformly distributed, so the joint distribution is uniform on $\prod_{s\in\mathcal{S}} \mathcal{Y}_s$. Let $\Lambda_\mathcal{S} = \Theta_\mathcal{S}(Y_s : s \in \mathcal{S})$, then $\Lambda_\mathcal{S}$ is uniform on $G/G_\mathcal{S}$ as $\Theta_\mathcal{S}$ is bijective. From Proposition 7, $\forall t \in \mathcal{S} \cup \mathcal{E}$, $Y_t = \pi'_t(\Lambda_\mathcal{S})$, and so by Lemma 6, the entropy vector for $\{Y_t : t \in \mathcal{S} \cup \mathcal{E}\}$ is characterizable by the group G and its subgroups $\{G_t : t \in \mathcal{S} \cup \mathcal{E}\}$.

C. Inclusion of Linear Network Codes

In this section we carry over the group theory notations in Section II to vector spaces, but with additive notation. For example, the left coset is now written as $v + W$ for a vector v and a subspace W. Further, we use $\oplus$ to denote the direct sum of vector spaces. In the following we show that for each linear network code, there exists an equivalent group network code, with essentially the same network operations and hence the same encoding/decoding results.

Consider a linear network code $\mathcal{C}$ over a finite field F. For each $t \in \mathcal{S} \cup \mathcal{E}$, the alphabet $\mathcal{Y}_t$ is a finite dimensional vector space over F. Let v denote the concatenation of all the source vectors $(Y_s : s \in \mathcal{S})$, then v is a vector in $V \triangleq \oplus_{s\in\mathcal{S}} U_s$, where $U_s \triangleq \mathcal{Y}_s$. Then for each edge e, the global mapping φ_e is a linear transformation from V to $\mathcal{Y}_e$, whose range is denoted by U_e. Also for each source s, let $\varphi_s : V \to U_s$ be the linear projection that maps $v \in V$ to its s-th section. Thus $\forall t \in \mathcal{S} \cup \mathcal{E}$, we can write $Y_t = \varphi_t(v)$. Let W_t be the null space of φ_t, then by the First Isomorphism Theorem,

$$\psi_t : v + W_t \mapsto \varphi_t(v)$$

is a vector space isomorphism between the quotient space V/W_t and U_t.

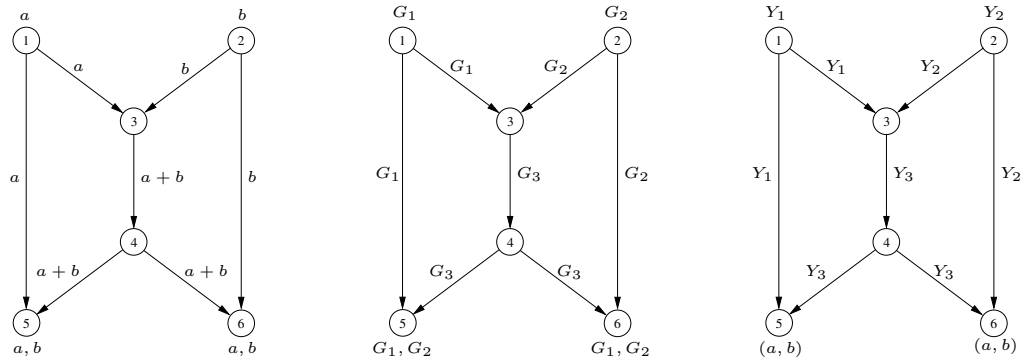
Let t be an edge or a sink node. If $Y_f = 0$ for all $f \in \mathcal{I}(t)$, then $Y_t = 0$ as the encoding/decoding functions are linear. Thus $\bigcap_{f\in\mathcal{I}(t)} W_f \leq W_t$. Further, for each source s

$$W_s = \{v \in V \mid s\text{-th section of v is } 0\} \cong \oplus_{r\in\mathcal{S}\setminus\{s\}} U_r,$$

so $\bigcap_{s\in\mathcal{S}} W_s = 0$. Since $V/W_s \cong U_s$, we have $\prod_{s\in\mathcal{S}} |V/W_s| = |V|$. Let $G = V$, $G_t = W_t$ for all $t \in \mathcal{S} \cup \mathcal{E}$. As V is a finite dimensional vector space over a finite field, G is a finite group. It is straightforward to check that the requirements (R1)–(R3) are all satisfied, so we can define a group network code $\mathcal{C}'$ with these groups.

This network code is equivalent to $\mathcal{C}$, since $\{\psi_t : t \in \mathcal{S} \cup \mathcal{E}\}$ provides a set of bijections between their codewords at each source/edge, and these bijections respect the encoding/decoding operations. In particular, assume in $\mathcal{C}$ the source vectors yield some $v \in V$, and so $Y_t = \varphi_t(v)$ is transmitted at each source/edge t. Then with ψ_t the corresponding symbol for $\mathcal{C}'$ is $v + W_t$, which is consistent with the encoding/decoding result of $\mathcal{C}'$ at each edge/sink node by Proposition 5.

For example, Fig. 3 demonstrates a linear network code over $\mathbb{F}_q$ for the well-known butterfly network (Fig. 3-(a)), and the corresponding group network code (Fig. 3-(b),(c)). Here for the linear network code, we have $V = \mathbb{F}_q^2$, $U_1 = U_2 = U_{e_{34}} = \mathbb{F}_q$, while $W_1 = \{(0, x) : x \in \mathbb{F}_q\}$, $W_2 = \{(y, 0) : y \in \mathbb{F}_q\}$, and $W_{e_{34}} = \{(z, -z) : z \in \mathbb{F}_q\}$. If we set $G = V$, $G_1 = W_1$, $G_2 = W_2$, and $G_3 = W_{e_{34}}$, then the resulting group network code is equivalent to the original linear one.



(a) (b) (c)

Fig. 3. Two network codes on the the butterfly network. (a) A linear network code; (b) the subgroup assignment for the corresponding group network code; (c) the transmitted symbols in the group network code. In (b), $G = \{(a, b) : a, b \in \mathbb{F}_q\}$, $G_1 = \{(0, x) : x \in \mathbb{F}_q\}$, $G_2 = \{(y, 0) : y \in \mathbb{F}_q\}$, and $G_3 = \{(z, -z) : z \in \mathbb{F}_q\}$. In (c), $Y_1 = \{(a, x) : x \in \mathbb{F}_q\}$, $Y_2 = \{(y, b) : y \in \mathbb{F}_q\}$, and $Y_3 = \{(a + z, b - z) : z \in \mathbb{F}_q\}$.

APPENDIX B
PROOFS AND CALCULATIONS IN SECTION VI

A. Structures of M, K, K', J, J'

When the characteristic p of $\mathbb{F}_q$ equals 2, $K = K'$ and $J = J'$. So for the analysis of K' and J' we only consider the case $p \neq 2$.

Observe that $|A_\alpha| = p$ for each $\alpha \in \mathbb{F}_q^\times$, and

$$|C| = 3, \quad |B_1| = 2, \quad |B| = |B'| = |P| = |P'| = q - 1.$$

As $(CB_1)^2 = I$, we have $M \cong D_6 \cong S_3$. It is easy to check that $\forall \alpha \in \mathbb{F}_q$,

$$A_\alpha^B = A_{t^{-1}\alpha}, \quad A_\alpha^{B'} = A_{-t^{-1}\alpha}, \quad A_\alpha^P = A_{t\alpha}, \quad A_\alpha^{P'} = A_{-t\alpha}.$$

Therefore, N is a normal subgroup of all K, K', J, J' and

$$K = N \cdot \langle B \rangle, \quad K' = N \cdot \langle B' \rangle, \quad J = N \cdot \langle P \rangle, \quad J' = N \cdot \langle P' \rangle.$$

Also N trivially intersects each of $\langle B \rangle, \langle B' \rangle, \langle P \rangle$ and $\langle P' \rangle$, thus

$$K \cong N \rtimes \langle B \rangle, \quad K' \cong N \rtimes \langle B' \rangle, \quad J \cong N \rtimes \langle P \rangle, \quad J' \cong N \rtimes \langle P' \rangle,$$

all of which are semidirect products $\mathbb{Z}_p^m \rtimes \mathbb{Z}_{q-1}$. We claim that $K \cong J$ and $K' \cong J'$. Moreover, in the case $p \neq 2$, all the four groups are isomorphic if and only if $\frac{q-1}{2}$ is even.

To see this, first consider the bijections $\sigma : K \to J$ and $\sigma' : K' \to J'$, where $\forall \alpha \in \mathbb{F}_q$, $\forall k \in \mathcal{K}_q$,

$$\sigma\left(A_\alpha B^k\right) = A_\alpha P^{-k}, \quad \sigma'\left(A_\alpha (B')^k\right) = A_\alpha (P')^{-k}.$$

Observe that $\forall \alpha, \beta \in \mathbb{F}_q$, $\forall k, l \in \mathcal{K}_q$,

$$\sigma\left(A_\alpha B^k \cdot A_\beta B^l\right) = \sigma\left(A_{\alpha + t^k \beta} B^{k+l}\right) = A_{\alpha + t^k \beta} P^{-k-l} = A_\alpha P^{-k} \cdot A_\beta P^{-l} = \sigma\left(A_\alpha B^k\right) \cdot \sigma\left(A_\beta B^l\right),$$

so σ is indeed an isomorphism. Similarly σ' is also an isomorphism.

Next observe that in the case $p \neq 2$, when $\frac{q-1}{2}$ is even, $\frac{q-1}{4}$ is an integer and so

$$\left(\frac{q+1}{2}\right)^2 = \left(\frac{q-1}{2} + 1\right)^2 = \frac{(q-1)^2}{4} + (q-1) + 1 \equiv 1 \pmod{q-1}.$$

Thus $\left((B')^{\frac{q+1}{2}}\right)^{\frac{q+1}{2}} = B'$ and $\left\langle (B')^{\frac{q+1}{2}} \right\rangle = \langle B' \rangle$. In addition, since $\mathbb{F}_q^\times$ is cyclic of an even order $q-1$, we have $-1 = t^{\frac{q-1}{2}}$, and thus $(-t)^{\frac{q+1}{2}} = \left(t^{\frac{q+1}{2}}\right)^{\frac{q+1}{2}} = t$. Consider $\tau : K \to K'$, where

$$\tau\left(A_\alpha B^k\right) = A_\alpha (B')^{\frac{q+1}{2}k}, \quad \forall \alpha \in \mathbb{F}_q, \ \forall k \in \mathcal{K}_q.$$

Apparently τ is a bijection. Also we can show that it is a homomorphism by calculating $\tau\left(A_\alpha B^k \cdot A_\beta B^l\right)$ with the following fact:

$$A_\alpha (B')^{\frac{q+1}{2}k} \cdot A_\beta (B')^{\frac{q+1}{2}l} = A_{\alpha + (-t)^{\frac{q+1}{2}k}\beta} (B')^{\frac{q+1}{2}(k+l)} = A_{\alpha + t^k \beta} (B')^{\frac{q+1}{2}(k+l)}.$$

Thus when $\frac{q-1}{2}$ is even, $K \cong K'$ and the four groups are all isomorphic.

When $\frac{q-1}{2}$ is odd, however, τ is not a bijection anymore, because this time $B' \notin \left\langle (B')^{\frac{q+1}{2}} \right\rangle$ and $\tau(K) \neq K'$. Furthermore, we can prove that in this case K and K' are not isomorphic, by showing that K and J have generalized flower structures whenever $q > 2$, whereas if $p \neq 2$, K' and J' only have flower structures when $\frac{q-1}{2}$ is even. Since $K \cong J$ and $K' \cong J'$, it is enough to only show the analysis of K and K'. Pick $\alpha \in \mathbb{F}_q^\times$ and assume $k, l \in \mathcal{K}_q$. Similar to the G_2 in Section V-B, we have the relation

$$(B^k)^{A_\alpha} = B^l \iff k = l = 0,$$

thus K has a generalized flower structure whenever $q > 2$. On the other hand, for K' we have

$$(B'^k)^{A_\alpha} = B'^l \iff \begin{bmatrix} (-1)^k & 0 \\ t^k \alpha & t^k \end{bmatrix} = \begin{bmatrix} (-1)^l & 0 \\ (-1)^l \alpha & t^l \end{bmatrix},$$

which requires $k = l$ and $t^l = (-1)^l$. Thus for $p \neq 2$, l can only be 0 or $\frac{q-1}{2}$. If $\frac{q-1}{2}$ is even, we have $(-1)^{\frac{q-1}{2}} = 1$ and so $k = l = 0$, then K' also has a generalized flower structure (as expected since here $K \cong K'$). If $\frac{q-1}{2}$ is odd, however, this is not true: in this case $(-1)^{\frac{q-1}{2}} = -1$, so $k = l = 0$ or $\frac{q-1}{2}$ in the above relation. Thus $\forall \alpha \in \mathbb{F}_q^\times$, $\langle B' \rangle \bigcap \langle B' \rangle^{A_\alpha} = \langle -I \rangle \cong \mathbb{Z}_2$. When $q = 3$, $B' = -I$ and

$K' = \langle A \rangle \times \langle -I \rangle \cong \mathbb{Z}_3 \times \mathbb{Z}_2 \cong \mathbb{Z}_6$; when $q > 3$, $\langle B' \rangle$ and $\langle B' \rangle^{A_\alpha}$ are distinct groups but have nontrivial intersection. Therefore, in neither case does K' have a generalized flower structure.

B. Intersections in Instances 8 and 9

Let $p \neq 2$. Observe that K' and J' are both subgroups of the G_2 in Instance 1, so all the intersections in both instances are subgroups of their respective counterparts in Instance 1. In Instance 8, since $G_{12} \leq \langle tI, B_1 \rangle$ and the (1,1)-entry for every matrix in $G_2 = K'$ is always ± 1, we have $G_{12} \leq \langle -I, B_1 \rangle$. This further limits the (2,2)-entry to be ± 1 for each matrix in G_{12}. As the (2,2)-entry in K' takes the form t^k for some k, this k can only be 0 or $\frac{q-1}{2}$. By examining the parity of $\frac{q-1}{2}$, we have

$$G_{12} = \begin{cases} \langle B_1 \rangle \cong \mathbb{Z}_2 & \text{if } \frac{q-1}{2} \text{ is even} \\ \langle -I \rangle \cong \mathbb{Z}_2 & \text{otherwise} \end{cases}, \quad G_{123} = G_{124} = \begin{cases} 1 & \text{if } \frac{q-1}{2} \text{ is even} \\ \langle -I \rangle \cong \mathbb{Z}_2 & \text{otherwise} \end{cases}.$$

Similarly we can calculate G_{12}, G_{123} and G_{124} for Instance 9.

In both instances, G_{24} is simply the subgroup of all diagonal matrices in G_2, and $G_{23} \leq T$. As matrices in K' and J' can be respectively written as

$$(-1)^k \begin{bmatrix} 1 & 0 \\ \alpha' & (-t)^k \end{bmatrix} = (-1)^k \begin{bmatrix} 1 & 0 \\ \alpha' & (t^{\frac{q+1}{2}})^k \end{bmatrix} \quad \text{and} \quad t^k \begin{bmatrix} 1 & 0 \\ \alpha'' & (-t^{-1})^k \end{bmatrix} = t^k \begin{bmatrix} 1 & 0 \\ \alpha'' & (t^{\frac{q-3}{2}})^k \end{bmatrix}$$

for some $\alpha', \alpha'' \in \mathbb{F}_q$ and $k \in \mathcal{K}_q$, we see that $G_{23} = \langle -B_3^{\frac{q+1}{2}} \rangle$ and $\langle tB_3^{\frac{q-3}{2}} \rangle$ respectively, where

$$(-B_3^{\frac{q+1}{2}})^k = \begin{bmatrix} (-1)^k & 0 \\ t^k - (-1)^k & t^k \end{bmatrix}, \quad (tB_3^{\frac{q-3}{2}})^k = \begin{bmatrix} t^k & 0 \\ (-1)^k - t^k & (-1)^k \end{bmatrix}.$$

Thus $G_{23} \cong \mathbb{Z}_{q-1}$ in both cases.

C. The case $p = 3$ for Instance 15

In Instance 15, $G_1 = M = \langle C, B_1 \rangle$ and $G_2 = (J')^E$. We can show that $G_1 G_2 = G_2 G_1$ when $p = 3$, thus Condition 3 is satisfied. Observe that $G_2 = \{X_{\alpha,j} \mid \alpha \in \mathbb{F}_q, j \in \mathcal{K}_q\}$, where

$$X_{\alpha,j} \triangleq \begin{bmatrix} (-1)^j - \alpha & \alpha \\ (-1)^j - t^j - \alpha & t^j + \alpha \end{bmatrix}.$$

When $p = 3$, we have $2 = -1$. With this relation, it is easy to check that $C = X_{1,0} \in G_2$, and for each α and j

$$X_{\alpha,j}^{B_1} = \begin{bmatrix} (-1)^j + \alpha & -\alpha \\ (-1)^j - t^j + \alpha & t^j - \alpha \end{bmatrix} = X_{-\alpha,j} \in G_2.$$

Thus G_1 normalizes G_2. In particular, $\forall X \in G_2$ and $\forall Y \in G_1$, we have $X^Y \in G_2$ and $X^{Y^{-1}} \in G_2$, which imply $XY \in G_1 G_2$ and $YX \in G_2 G_1$ respectively. Therefore $G_1 G_2 = G_2 G_1$.

D. Intersections in Instances 12–15

Most intersections are easily obtained by comparing the formulae of the matrices in the subgroups involved. For the intersection of M with any of $J^E, (J')^E, J^Q$ or $(J')^Q$, we can utilize the properties below to facilitate calculation. Let $\vec{c}_i(X)$ denote the i-th column of a matrix X, we have

$$\vec{c}_1(X)+\vec{c}_2(X)=\begin{bmatrix}1\\1\end{bmatrix},\ \forall X\in J^E;\qquad \vec{c}_1(X)+\vec{c}_2(X)=\pm\begin{bmatrix}1\\1\end{bmatrix},\ \forall X\in (J')^E;$$

$$\vec{c}_1(X)-2\vec{c}_2(X)=\begin{bmatrix}1\\-2\end{bmatrix},\ \forall X\in J^Q;\qquad \vec{c}_1(X)-2\vec{c}_2(X)=\pm\begin{bmatrix}1\\-2\end{bmatrix},\ \forall X\in (J')^Q.$$

Thus, we need only seek elements of M which share these properties.

We also want to mention the calculation of G_{34} for Instances 13 and 15 when $p>3$. In Instance 13, finding G_{34} is equivalent to solving the following set of equations:

$$\begin{cases}(-1)^j-\alpha=(-1)^i+2\beta\\ \alpha=\beta\\ (-1)^j-t^j-\alpha=2\left(t^i-2\beta-(-1)^i\right)\\ t^j+\alpha=t^i-2\beta\end{cases}\iff\begin{cases}\alpha=\beta\\ 3\beta=(-1)^j-(-1)^i\\ t^i=(-1)^j\\ t^j=(-1)^i\end{cases}.$$

From the last two equations, we can see that i and j can only be 0 or $\frac{q-1}{2}$. If $\frac{q-1}{2}$ is even, then $(-1)^{\frac{q-1}{2}}=1$, so i and j must both be 0, which yields that $G_{34}=1$. If $\frac{q-1}{2}$ is odd, then $i=0$ implies that $j=0$, and $i=\frac{q-1}{2}$ implies that $j=\frac{q-1}{2}$. In both cases $\alpha=\beta=0$, therefore $G_{34}=\langle -I\rangle$. For G_{34} in Instance 15, we have similar equations and the same discussion also applies.

ACKNOWLEDGMENT

The authors would like to thank Michael Aschbacher and Amin Shokrollahi for very helpful discussions on the conditions and the generalizations of the violations, and on expanding the group structures. They would also like to thank Ryan Kinser for reminding them of the case $PGL(2,q)$.

REFERENCES

[1] W. Mao and B. Hassibi, "Violating the Ingleton inequality with finite groups," in *Proc. of the 47th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Sep./Oct. 2009.

[2] W. Mao, M. Thill, and B. Hassibi, "On group network codes: Ingleton-bound violations and independent sources," in *Proc. of the 2010 IEEE International Symposium on Information Theory*, Austin, TX, Jun. 2010.

[3] T. H. Chan and R. W. Yeung, "On a relation between information inequalities and group theory," *IEEE Trans. Inf. Theory*, vol. 48, no. 7, pp. 1992–1995, Jul. 2002.

[4] R. Dougherty, C. Freiling, and K. Zeger, "Insufficiency of linear network coding in network information flow," in *IEEE Transactions on Information Theory*, 2005, pp. 2745–2759.

[5] Z. Zhang and R. W. Yeung, "A non-Shannon-type conditional inequality of information quantities," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1982–1986, Nov. 1997.

[6] X. Yan, R. Yeung, and Z. Zhang, "The capacity for multi-source multi-sink network coding," in *Proc. of 2007 IEEE International Symposium on Information Theory*, Nice, France, Jun. 2007, pp. 116–120.

[7] B. Hassibi and S. Shadbakht, "Normalized entropy vectors, network information theory and convex optimization," in *Proc. of the 2007 IEEE Information Theory Workshop*, Jul. 2007, pp. 1–6.

[8] R. Dougherty, C. Freiling, and K. Zeger, "Networks, matroids, and non-shannon information inequalities," in *IEEE Transactions on Information Theory*, June 2007, pp. 1949–1969.

[9] S. Shadbakht and B. Hassibi, "Cayley's hyperdeterminant, the principal minors of a symmetric matrix and the antropy region of 4 Gaussian random variables," in *Proc. of the 46th annual Allerton Conference on Communication, Control and Computing*, Sep. 2008.

[10] T. Chan, "A combinatorial approach to information inequalities," in *Commun. Inf. Syst.*, vol. 1, no. 3, 2001, pp. 241–253.

[11] A. Ingleton, "Representation of matroids," in *Combinatorial Mathematics and its Applications*, 1971, pp. 149–167.

[12] D. Hammer, A. Romashchenko, A. Shen, and N. Vereshchagin, "Inequalities for shannon entropy and kolmogorov complexity," *Journal of Computer and System Sciences*, vol. 60, no. 2, pp. 442–464, Apr. 2000.

[13] F. Matúš, "Conditional independences among four random variables III: Final conclusion," *Combinatorics, Probability and Computing*, vol. 8, pp. 269–276, 1999.

[14] R. Kinser, "New inequalities for subspace arrangements," *Journal of Combinatorial Theory*, vol. 118, no. 1, pp. 152–161, Jan. 2011.

[15] R. Dougherty, C. Freiling, and K. Zeger, "Linear rank inequalities on five or more variables," 2010, preprint. [Online]. Available: arXiv:0910.0284v3

[16] R. Dougherty, "Computations of linear rank inequalities on six variables," in *Proc. of the 2014 IEEE International Symposium on Information Theory*, Jun. 2014, pp. 2819–2823.

[17] R. Dougherty, C. Freiling, and K. Zeger, "Characteristic-dependent linear rank inequalities and network coding applications," in *Proc. of the 2014 IEEE International Symposium on Information Theory*, Jun. 2014, pp. 101–105.

[18] T. Chan, A. Grant, and D. Pflüger, "Truncation technique for characterizing linear polymatroids," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6364–6378, Oct. 2011.

[19] T. H. Chan, "Group characterizable entropy functions," in *Proc. of the 2007 IEEE International Symposium on Information Theory*, Nice, France, Jun. 2007, pp. 506–510.

[20] ——, "On the optimality of group network codes," in *Proc. of the 2005 IEEE International Symposium on Information Theory*, Adelaide, Australia, Sep. 2005, pp. 1992–1996.

[21] ——, "Capacity regions for linear and abelian network codes," in *Proc. of the 2007 Information Theory and Applications Workshop*, La Jolla, CA, Jan./Feb. 2007, pp. 73–78.

[22] Z. Zhang and R. W. Yeung, "On characterization of entropy function via information inequalities," *IEEE Trans. Inf. Theory*, vol. 44, no. 7, pp. 1440–1452, Jul. 1998.

[23] N. Boston and T.-T. Nan, "Large violations of the Ingleton inequality," in *Proc. of the 50th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Oct. 2012.

[24] P. Paajanen, "Finite p-groups, entropy vectors, and the Ingleton inequality for nilpotent groups," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 3821–3824, Jul. 2014.

[25] N. Markin, E. Thomas, and F. Oggier, "Groups and information inequalities in 5 variables," in *Proc. of the 51th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Oct. 2013.

[26] D. S. Dummit and R. M. Foote, *Abstract algebra*, 3rd ed. Hoboken, NJ: Wiley, 2004.

[27] *GAP – Groups, Algorithms, and Programming, Version 4.4.12*, The GAP Group, 2008. [Online]. Available: http://www.gap-system.org

[28] H. Li and E. K. P. Chong, "On connections between group homomorphisms and the ingleton inequality," in *Proc. of the 2007 IEEE International Symposium on Information Theory*, Nice, France, Jun. 2007, pp. 1996–2000.

[29] D. L. Johnson, *Presentations of Groups*, ser. London Mathematical Society Student Texts. Cambridge: Cambridge University Press, 1990, vol. 15.

[30] R. Dougherty, C. Freiling, and K. Zeger, "Non-shannon information inequalities in four random variables," 2011, preprint. [Online]. Available: arXiv:1104.3602v1

[31] F. Matúš and L. Csirmaz, "Entropy region and convolution," *Combinatorics, Probability and Computing*, submitted. [Online]. Available: arXiv:1310.5957v1

[32] S. Shadbakht and B. Hassibi, "MCMC methods for entropy optimization and nonlinear network coding," in *Proc. of the 2010 IEEE International Symposium on Information Theory*, Austin, TX, Jun. 2010.

[33] S. Shadbakht, "Entropy region and network information theory," Ph.D. dissertation, California Institute of Technology, 2011.